

September, 1961

Vol. 32, No. 3

Child Development

REC'D SEP 5 - 1961

CONTENTS

A Human Biologist Looks at the Contemporary Scene Wilton Marion Krogman	417
An Investigation of Maternal Concerns by Mail Survey William O. Robertson	423
Positive Reinforcement and Behavioral Deficits of Autistic Children C. B. Ferster	437
Comparison of Data Obtained from Mothers and Fathers on Childrearing Practices and Their Relation to Child Aggression Leonard D. Eron, Thomas J Banta, Leopold O. Walder, and Jerome H. Lullicht	457
Q-Sort Descriptions of Five-Year-Old Children by Their Parents Gene R. Medinuss	473
The Reliability of Mothers' Histories Charles Wenar	491
Family Relations of Bright High-Achieving and Under-Achieving High School Boys William R. Morrow and Robert C. Wilson	501
Peer Popularity and Dependence on Adults in Preschool-Age Socialization Boyd R. McCandless, Carolyn B. Bilous, and Hannah Lou Bennett	511
Sex of Experimenter and Sex-Typed Behavior of Young Children L. J. Borstelmann	519
Maternal Acceptance of Retarded Children: A Questionnaire Study of Attitudes and Religious Background G. H. Zuk, Ralph L. Miller, John B. Bartram, and Frederick Kling	525
Recognition of Perceptually Ambiguous Stimuli in Grade School Children Juris G. Draguns and Giuseppe Multari	541
Quantity Conceptions in Junior and Senior High School Students David Elkind	551
Studies of Disability: A Technique for Measurement of Psychological Effects Robert W. Downing, George Moed, and Byron W. Wight	561
Response Sets and Manifest Anxiety Scores in a Retarded Population Charles G. Matthews and Leon H. Levy	577
Relations between Nursery School Attendance and Teachers' Ratings of Some Aspects of Children's Adjustment in Kindergarten Ann Wilson Brown and Raymond G. Hunt	585
Prediction of Schoolage Intelligence from Infant Tests Sibylle K. Escalona and Alice Moriarty	597
Comment on "The Relationship between Rigidity-Flexibility in Children and Their Parents" John William Asher	607

Published Quarterly by Child Development Publications of the

SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT, INC.

Editor

WILLIAM E. MARTIN

Purdue University

Consulting Editors

DONALD M. BAER

University of Washington

EMANUEL K. BELLER

Child Development Center

ALFRED CASTANEDA

University of Texas

WALTER EMMERICH

University of Colorado

JEROME KAGAN

Fels Research Institute

IRVIN A. KRAFT

Houston State Psychiatric Institute

LEWIS P. LIPSITT

Brown University

NORMAN LIVSON

University of California

JAMES M. TANNER

University of London

RICHARD H. WALTERS

University of Toronto

Managing Editor

KATE HOFFMAN

Child Development, one of three publications of the Society for Research in Child Development, Inc., is issued quarterly in March, June, September, and December. The subscription price per year is \$12.00 domestic, \$12.50 foreign. Single issues are \$3.50.

Child Development Abstracts and Bibliography is issued three times a year, two numbers in each issue. The subscription price per year is \$6.00 domestic, \$6.50 foreign. Single issues are \$2.50.

Monographs of the Society for Research in Child Development is issued at irregular intervals during the year. The subscription price per year is \$6.00 domestic, \$6.50 foreign. The price of individual monographs is variable.

The three publications of the Society are available at the special price of \$20.00 per year, domestic; \$21.50 per year, foreign.

A limited number of back issues of all publications is available. A mimeographed list may be had upon request.

Subscriptions, address changes, and business communications should be sent to CHILD DEVELOPMENT PUBLICATIONS, Purdue University, Lafayette, Indiana.

Please give four weeks' notice when changing your address, giving old and new addresses. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Printed and mailed by
THE ANTIOCH PRESS
Yellow Springs, Ohio



A HUMAN BIOLOGIST LOOKS AT THE CONTEMPORARY SCENE¹

WILTON MARION KROGMAN*

University of Pennsylvania

At the risk of overemphasis I'd like to observe that biology is the fact of life, culture the factor of living; biology is the substance of life, culture the form of living. Biology is thus the very nucleus and matrix of all that is. All else is peripheral. Biology provides the force and direction to the stream of life. All else provides environmental background of living and existence.

As a human biologist I'd like to observe that we cannot evaluate the contemporary scene as an achieved isolate. What *is* happening to Man *has* happened to Man. What we contemplate Today has its roots in Yesterday. Man is an evolving, not an evolved, creature, and, I am constrained to say, this is in principle even more true of what he has wrought in terms of cultural complexities. We reap a whirlwind Now because a wind was sown Then!

What I propose to do it to limn our problems in the perspective of evolutionary time, for Man is one end-result of eons of emergent development. We begin, simply, with a ganglion—an aggregation of nerve cells—at the anterior end of an ancient form of life. This ganglion, this raw knot of action and response cells, was the forerunner of the *brain* which was to become the guiding agent first in more or less random response, then in coordinated reflex interaction, and finally in conscious and directed behavioral reactions dictated by a cortex that had evolved into a *mind*. Now the brain no longer runs a body alone, it directs and guides a life! From mere neurophysiology has arisen the incredible and tortuous complexity of *social behavior*. To the organic has been added the supraorganic!

Out of all of the many facets of the current cultural scene I'd like to focus, briefly, upon two: human relationships and human numbers, both, in my opinion, biologically determined, but culturally oriented and mediated.

* Graduate School of Medicine, University of Pennsylvania, Philadelphia.

¹ Address of the President at the Biennial Meeting of the Society for Research in Child Development, Pennsylvania State University, March 17, 1961.

Child Developm., 1961, 32, 417-422.

CHILD DEVELOPMENT

I refer, of course, to problems of race and to problems of population composition. I shall consider them in this order.

There is a major premise to be accepted at once: all people now living, or who have lived for the past 100,000 years or so, belong to the same genus, *Homo*, and to the same species, *sapiens*. To make the identity even more graphic, it may be observed that 29,500 pairs of the 30,000 pairs of human genes are shared equally by all of Mankind. It is the remaining 500, withdrawn differentially from the common gene pool, that set up so-called "racial differences"—those external hallmarks of skin, hair, nose that grant measures of unlikeness. In the broadest possible terms there is but one "human race."

Notwithstanding the foregoing generalization, there *are* subdivisions called races. What, then, is a race? Genetically speaking, a race is a breeding population within *Homo sapiens* with a certain combination of heritable traits to give a degree of physical distinctiveness. The population unit is subject to a certain measure of isolation, either geographical (formerly) or culturally (currently).

How many races are there? It depends upon the criteria, or combination thereof, set up for classification. There are "macrogeographical" and "microgeographical" races. Hence, the number may vary from three or four to 30. Today we generally accept eight major geographical races: European, African, Asiatic, (East) Indian, Australian, Melanesian, Polynesian, and Amerindian.

Our concern is not really with either definition or classification of race: it is with cultural *reaction* to the biological fact of race. The problem is not one of whether or not a person looks differently, but the *assumption of difference*. There are four pretty well-defined reactions to and beliefs about, race differences: (a) they are fixed, immutable, unassimilable; (b) so-called racial traits (behavioral, intellectual) are innate, i.e., genetically entrenched; (c) the greater the degree of *seen* difference, the greater the separation in a superior-inferior hierarchy; (d) hybridization ("miscegenation") can only result in biological inferiority or deterioration ("the lower will drag the upper down"). These are the ghosts which must be laid before "race problems" cease to haunt us.

In 1950 the population of the United States was composed as follows:

Native white	124,780,860
Foreign-born white	10,161,168
Negro	15,042,286
Indian	343,610
Japanese	141,768
Chinese	117,629
Other	110,240

In a sense this breakdown is really ludicrous. There are no "native" whites; the only true "natives" are the Amerindians. Be that as it may,

we Americans are neither biologically nor culturally homogeneous. There are differences among us: in appearance, in cultural level, in performance, in behavior and attitude, in achievement, and in ability and potential. These are an observable and measurable set of facts. The problem is whether or not these variable levels are race-linked. Is there a variability, in a given direction, that is innate? Is variability interracial, or is it intraracial? For racial friction do we blame *Biology* or do we penalize *Culture*? What is behind a 1944 report that we Americans ranked peoples in order of most to least preferred: white American, Canadian, English, French and Swiss, Spanish and Portuguese, Jew, Negro, Turk, Chinese? Why such a cultural hierarchy?

In purely biological terms there is no validity to claims of "superior" racial status. As we move from a solely structural consideration of race, based upon morphology, to a functional analysis of race, based upon physiology and biochemistry, we are beginning to realize that there are certain adaptive traits that possibly confer greater survival values, with specific reference to the ecological niche occupied at the time the type was evolving (being selected). That is to say, for example, that a pigmented skin is better adapted (hence, "superior") to a tropical sun; or a prominent malar complex with fat pads to a sub-Arctic or Arctic climate.

It is time to quit calling a spade an agricultural implement. It is time to come to grips not with race, but with *racism*. It is time to face up to the fact of *prejudice*: the creation of advantages and material benefits for the prejudicor, the limitation or denial of these to the prejudictee.

We are making progress. The FEPC, antidiscrimination decrees in the armed forces, and desegregation in the schools are all ameliorative. But we continue to pay a high sociocultural price for barriers of prejudice. In our economy prejudice has caused an unbalanced use of manpower which, in turn, has set up a maze of welfare problems. In broader perspective prejudice has created a national and political schism and aggravated regional tensions. Prejudice is a hindrance to free and clear sociopolitical communication, both at national and international level. It engenders a frustration-aggression pattern of social interaction, and it warps the attitude and personality (group and individual) of those prejudiced. Prejudice feeds on itself and is contagious: it is an easy progression from anti-Negro to anti-Jew, or anti-Catholic, or anti-upper class, or what you will.

I repeat, we are making progress. But, I have heard it said, "knowledge is slow of foot, and wisdom limps far behind knowledge." In current terms we are witnessing a lag between legislative and judicial action and socio-cultural implementation. In oversimplified terms the issue of school desegregation is a conflict between law and custom. Since law represents the codified will of the majority, there can be no doubt of either its validity or its ultimate and absolute triumph.

I think that we are all coming to realize that problems of race relations are generic: the Negro-white problem is really one with the minority-domi-

CHILD DEVELOPMENT

nant problems of Catholics, Jews, Spanish-Americans, Japanese-Americans, and so on. I feel strongly that we Americans can resolve our so-called race problem only if we come to the realization that such a problem is little or no different from any other intergroup or interpersonal problem in our everyday cultural life. I have every hope that experience, reason, understanding, and a give-and-take attitude will triumph over prejudice and its accompanying complex of ignorance, fear, and emotion.

Here is where we, people of goodwill working with children, can do a lot. It is said that race prejudice must be learned. We can, by precept and by admonition, help to see that it is either not learned or unlearned. The receptive mind of the child will store ideas, concepts, and examples of "fair play," of equality of opportunity, of sharing the good things of the American way of life. Democracy, full and even participation of the everydayness of living, is not a race-limited or a part-time affair.

I want to turn now to an even more critical problem. It, too, is biological: what *can* be more biological than breeding? I refer, of course, to the aptly-termed "population explosion."

Once more I hold to the long-term view. In his earlier days *Homo* was irregularly and sparsely spread over Eurasia. Not until the end of the Ice Age, some 25,000 years ago, did he begin his deployment. Since that time two great cultural events have spurred his increase and his spread: (a) the "Agricultural Revolution" of about 10,000 years ago; (b) the "Industrial Revolution" (or "Scientific Revolution") of about 1650 A.D. Man in the Old Stone Age may have numbered 10 million; in 6000 B.C. he could count 20 million; by 1200 B.C., 100 million; at the opening of the Christian Era, 350 million; by 1650 A.D. varying estimates run from 470 million to one billion. (In Europe, alone, the 1650 to 1950 A.D. increase was from 100 to 900 million).

In more realistic terms we may summarize as follows: Man achieved his first billion by 1830; the second by 1930 (100 years elapse); his third by 1965 (35 years elapse); his fourth by 1980 (15 years elapse); his fifth by 1990 (10 years elapse); his sixth by 2000 (10 years elapse). To drive this closer home, the United States will add 100 million by 1980, to bring us to 280 million. One of every 20 men who have ever lived is alive today. The time is not too far away when a tenth of all humans who have ever lived will walk the Earth simultaneously! If your pulse ran 100 to the minute, every pulse beat would record a birth!

In the United States we are facing a problem of urbanization. In 1950 about 57 per cent of the United States population lived in 168 metropolitan areas (cities of 50,000 plus). From 1900 to 1950 such areas absorbed 73 per cent of total population growth; in the last decade of this period the percentage rose to 81. By 1975 we shall have 200 such metropolitan areas.

Over and beyond the urban trend is the likelihood of uneven geographic distribution in the United States. By areas the percentage growth for 1955 to 1970 may be calculated as follows: (a) 51 per cent for the three Pacific

States; (b) 44 per cent for the eight Mountain States; (c) 32 per cent for the five East-North-Central States; (d) 28 per cent for the eight South Atlantic States; (e) 22 per cent for the three Mid-Atlantic States; (f) 21 per cent for the four West-South-Central States; (g) 16 per cent for the seven West-North-Central States; (h) 16 per cent for the six New England States; and (i) 5 per cent for the four East-South-Central States.

Still being ethnocentric and looking at the United States, what are some of the consequences? Let me cite a few quite at random. The 65 million cars now on the highways if placed bumper-to-bumper 10 abreast on a 10-lane highway would encircle the Equator. Everytime we build a mile of superhighway we lose 40 acres of land. In a little over 10 years we will have dumped 60 million gallons of radioactive residue, with a life-span of from a decade to 24,000 years. Millions of Americans now drink treated sewage. In 1919 to 1956 the United States population increased 73 per cent, but the number on the government payroll went up 166 per cent. In 1912 a Senator represented 993,000 people, a Congressman 219,000; in 1960 the numbers were, respectively, 1,800,000 and 400,000; by 1975 they will be 2,225,000 and 517,000. By 1965 one half of all children born in New York City will be born to indigent families. I could go on and on with data such as these; some of them almost seem to be *non sequiturs*, but they are all tied in, one way or another, with more and more people—and still more!

In an over-all view the problem of numbers resolves to two considerations: *reduce them or feed them*. The first makes more sense to me, with a full realization of the ethical and moral issues. Controlled, i.e., limited, breeding is the only solution: world-wide contraception. The real issue centers around the ultimate morality of underprivileged conditions for one half of a tremendous population or more evenly distributed cultural and material benefits for a smaller population. I think we must squarely face the issue: have we the moral right to subordinate *quality* of population to *quantity*? It reduces to this: do we favor a lot of "have-nots" or fewer "haves"? When I refer to quality vs. quantity, have-not vs. have, I am referring to *both* biology and culture. I include both materialistic and spiritual values. The squalor and debasing circumstances of rank overpopulation must leave their imprint upon body and spirit alike. The dignity of complete individuality must not be swallowed in the morass of overwhelming numbers.

The land area of the Earth is about 55 million square miles, or about 36 billion acres; only about 2 per cent is readily arable (though Man has stretched this to 10 per cent). If our nearly three billion population were evenly distributed, there would be 40 people per square mile. Yet 66 per cent of the three billion live on only 8 per cent of the land surface. The end result? Anywhere from 66 to 75 per cent of the three billion live on a subminimal diet; they go to bed hungry every night. If we are to remedy this, the food supply will have to be doubled within the next generation

CHILD DEVELOPMENT

(say, 25 years). Answers are available, in potential at least: supplementary irrigation; more extensive use of fertilizers; plant breeding for specific environment and for insect and disease resistance; the use of special chemical media, hydroponics, for example. If the foregoing could be completely and integratively developed, we could hope for a 25-fold yield increase, which could, in theory, support a total population of 50 billion.

There is a very important corollary of *absolute* numbers, and that is the problem of *relative* numbers. Not only are more people living, but they are living longer! By 1800 A.D., after nearly 2000 years of post-Roman history, Man achieved a life-expectancy of 35 years. From 1800 to 1950 this was *doubled* to 70 years. This is to say, we have made more gain in the past 150 years than in all recorded and prerecorded human history.

Man now is at a stage where he, alone of all evolved life, may direct and influence his future evolution. It is not so much of the body that I speak, for I feel it will not change too much in the next 100,000 years or so (unless, of course, radiation does open the Pandora's box of unpredictable mutations). It is the mind, the spirit, the very soul, of Man that will be focal in his self-mediated evolutionary path. Man's value-needs will guide the emergent human being. It has been said that Man must have a "feeling of certitude," that the Darwinian "struggle for existence" has given way to the (perhaps Freudian?) "struggle for fulfillment." In essence this cuts across all human problems of race, of numbers, of age-levels, for here is implicit the essential value of the human individual. At this level fulfillment means the facilitation of his own personality, the realization of his own innate and learned capacities, the satisfaction of material and spiritual needs and drives, and the development throughout life of new qualities of experience. Here, for each human being, there should be no barriers of race, color, creed, socioeconomic status, or age-level!

I envision Man—all men—as sharing equally the greatness of what *Homo sapiens* has wrought: Man makes conscious choices, governed by a conscience; Man works with abstractions and conceptual frameworks, so that right and wrong, in both abstract and concrete, may be adjudged; Man formulates his own ultimate goals and establishes the values to be achieved; Man has defined and created beauty for himself and set in motion the mechanisms for its realization and expression; and Man has formulated the moral and ethical qualities of evil vs. good, with guilt the arbiter or penalty of choice or evaluation—Man's the freedom to choose and to pursue.

I conclude what I have to say with an affirmation of faith not alone in Man but in God. If in a sense Man has walked through the ages in accordance with natural law, he will continue to do so. Yet, I would rephrase it in the words of the Psalmist of old who admonished that Man "walk humbly" with his God. In this sense it is Man's destiny to stride confidently into the future with a renewed and deepened faith in the God of all life.

AN INVESTIGATION OF MATERNAL CONCERNS BY MAIL SURVEY

WILLIAM O. ROBERTSON*

The Children's Hospital, Columbus, Ohio

In recent years, pediatric attention has been increasingly directed toward well child supervision. Prevention rather than treatment of "disease" is the ultimate objective. An important facet of the program that has evolved is the pediatrician's function in advising and reassuring mothers about the common concerns of child rearing.

Surprisingly little is to be found in the literature about the focus or incidence of these common concerns. Admittedly, many factors influence the development of maternal concern, its transmission to the physician, and his perception and interpretation of it. The complexity of these influences and the costs of investigations that might control each variable effectively make it understandable that a paucity of objective data on the subject exists. Nevertheless, if the "New Pediatrics" (5) is to progress, reliable norms about this as well as other aspects of maternal behavior would appear essential.

The mail questionnaire method of survey has established its value in various nonmedical fields. Currently it is also being employed effectively in medical research; for example, in longitudinal studies on the incidence of cigarette smoking (10) and follow-up studies on the dissemination of pathogenic staphylococci from newborn nurseries (9). Consequently, it seemed worthwhile to explore the usefulness of this technique in securing childrearing data. Therefore, a mail questionnaire survey directed to mothers of young infants was undertaken in March, 1957. To determine, among other things, the consistency of its findings, it was repeated in November, 1958.

METHODOLOGY

A mail questionnaire listing various sources of potential maternal concern was developed. Explanatory instructions requested that the mother examine each category and note her answer simply by checking the appropriate response. Space was provided for her to describe or explain the nature of any existing disease or abnormality, as well as her physician's recommended

* The Children's Hospital, Columbus 5, Ohio.

CHILD DEVELOPMENT

therapy, and to add comment she might consider appropriate. Certain specific items, such as the infant's sex, the maternal age, etc., were also requested, to provide data for distributional analysis and correlative purposes.

Preliminary trials were undertaken to test adequacy of instructions, comprehension of terminology, and degree of response. Necessary modifications were then incorporated into the questionnaire and its final format prepared, as shown in the Appendix.

The names of 2,000 mothers who had delivered during a specified interval were obtained from existing commercial sources. These names are alleged to have been assembled in a random fashion from a master list said to include 85 to 90 per cent of the nation's births. Analysis of their geographical distribution showed them to be statistically representative.

A letter, signed by the author, was then forwarded, together with a questionnaire, to each name listed. It congratulated the mother and requested that she complete and return, anonymously, the enclosed questionnaire in an accompanying prepaid addressed envelope. One week later, a second mailing of the letter, questionnaire, and envelope, supplemented with a notation thanking the mother if she had already returned the questionnaire, and requesting her to do so if she had not, was dispatched.

Replies on returned questionnaires were coded and entered on standard IBM cards. The geographical location of each respondent was determined from the postal marking on the returned envelope. These locations were assigned their specific county sites and appropriate population figures determined from available government statistics (12). Arbitrarily, those counties whose populations exceeded 100,000 or which bordered large metropolitan centers were designated as "urban"; otherwise they were "rural." Appropriate designations of "urban" and "rural" were entered on the individual IBM cards.

The distribution of the group responding was compared to control data with regard to regional geographical location, age, education, and parity of the mother, as well as sex, age, and birth weight of the infant (11, 12, 13). Incidences of different concerns were computed for specified variables.

Response

In this pilot investigation, two independent surveys were conducted, utilizing an identical questionnaire. The first, Survey A, was initiated in March, 1957. Of 2,000 mailed questionnaires 125 could not be delivered because of the addressee's temporary or permanent change of address or incomplete address. These were returned unopened. Of the 1,875 delivered questionnaires, 1,132 (60.5 per cent) were completed and returned. Thirty-nine were discarded because of (a) grossly inadequate completion of the form, (b) illegible postmarking on the return envelope, or (c) late completion or return of the questionnaire. A total of 1,093 returns was coded and entered on IBM cards for processing. The second survey, Survey B,

TABLE I
INCIDENCE OF MATERNAL CONCERN

General Category	PERCENTAGE OF MOTHERS INDICATING				TOTAL WORRIES		
	"Some Worry"		"Considerable Worry"				
	Survey A	Survey B	Survey A	Survey B	Survey A	Survey B	Survey A & B
1. <i>Digestion</i> (spitting, burping, gas, vomiting, colic, etc.)	34.7	33.7	7.7	7.6	42.3	41.4	41.9
2. <i>Breathing</i> (uneven, hic-coughs, gags, chokes, gasps, grunts, etc.)	34.8	32.2	4.2	3.8	39.0	36.0	37.5
3. <i>Bowel Movements</i> (odor, color, too often, too loose, hard, etc.)	30.9	29.0	7.4	7.4	38.2	36.4	37.3
4. <i>Skin</i> (oily, dry, rash, scratches, etc.)	29.2	28.2	3.7	3.3	32.9	31.4	32.1
5. <i>Sleeping</i> (too much, not enough, not regular, restless, etc.)	27.1	24.9	4.9	4.2	32.0	29.1	30.6
6. <i>Eating</i> (not enough, too much, not regular, hun-gry, disagrees, etc.)	25.3	22.5	5.2	4.0	30.5	26.5	28.5*
7. <i>Navel</i> (swollen, too large, small, bleeding, odor, etc.)	23.3	19.6	4.8	5.1	28.1	24.7	26.4
8. <i>Buttocks</i> (diaper rash, sore, color, etc.)	20.0	18.8	3.0	4.7	23.0	23.5	23.2
9. <i>Other</i> (spoiling baby, food preparation, bathing, clothing, diapering, etc.)	22.0	19.0	2.7	2.4	24.7	21.3	23.0
10. <i>Crying</i> (too much, too little, strong, weak, turn color, etc.)	17.3	14.7	4.3	3.4	21.7	18.1	19.9*
11. <i>Eyes</i> (puffy, red, crossed, color, etc.)	17.3	14.6	1.8	2.1	19.1	16.7	17.9
12. <i>Accidents</i> (while sleep-ing, eating, playing, bathing, etc.)	15.9	11.9	2.4	2.4	18.3	14.3	16.3*
13. <i>Head</i> (size, shape, soft spot, etc.)	12.5	10.9	2.5	1.7	15.0	12.6	13.8
14. <i>Mouth-Lips</i> (size, shape, color, sore, swallowing, thumbsucking, etc.)	7.8	9.5	1.4	1.8	9.1	11.2	10.1
15. <i>Weight</i> (not gaining, too fat, too thin, etc.)	7.1	6.9	0.6	1.7	7.6	8.6	8.1
16. <i>Nose</i> (size, shape, running, etc.)	8.4	6.0	0.6	1.0	8.9	7.0	8.0
17. <i>Legs-Feet</i> (too thin, too heavy, not straight, etc.)	5.5	5.1	1.7	1.2	7.2	6.4	6.8
18. <i>Urine</i> (odor, color, too often, too little, etc.) . . .	4.4	6.9	0.6	.8	5.0	7.7	6.3*
19. <i>Ears</i> (shape, size, color, etc.)	4.7	5.1	0.5	.7	5.2	5.8	5.5
20. <i>Stomach</i> (too large, small, hard, soft, swollen, etc.)	4.5	4.0	0.5	.7	5.0	4.7	4.8
21. <i>Hair</i> (too much, too little, falling out, etc.)	3.5	3.7	0.6	.2	4.1	3.9	4.0
22. <i>Arms-Hands</i> (too thin, too heavy, not straight, etc.)	1.0	1.2	0.2	.3	1.2	1.4	1.3

* Differences between results of Survey A and B significant at .05 level but not .01 level.

CHILD DEVELOPMENT

was undertaken in November, 1958. Of 2,500 questionnaires distributed, 2,398 were delivered; 1,134 (47.5 per cent) were completed, returned, and analyzed.

When experimental data on age, parity, and level of education of the mothers, sex and birth weight of the infants, and geographical location of the families were compared with known data, no statistically significant differences were found except that, as was anticipated, the level of education of the respondents exceeded that found among women of comparable age groups in the general population. The age of the infants at the time the mothers completed the questionnaires ranged from 5 to 10 weeks; the mean age was 7 weeks.

RESULTS

The frequencies with which different sources of concern were noted in the two surveys are listed in Table 1. Comparison of their findings shows that the incidence of "total worries" is significantly different at the 5 per

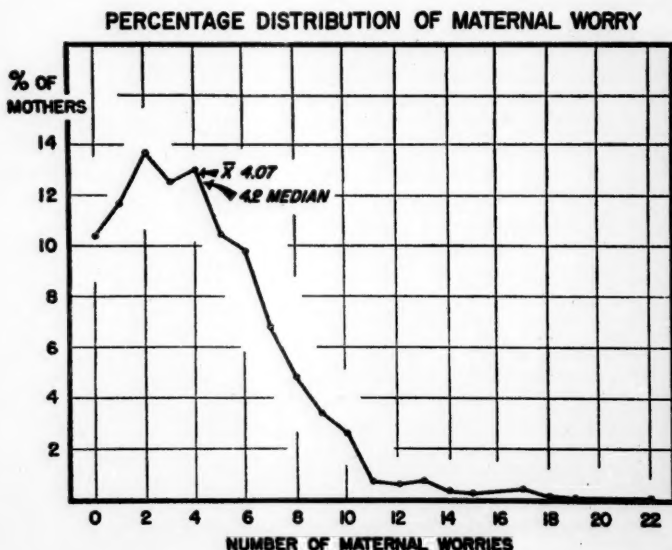


FIGURE 1—A distribution curve depicting the percentage of respondents checking none through the 22 items listed in the questionnaire. The mean number of items noted is 4.1; the median is 4.2. Of the mothers 10.4 per cent failed to note any item as a source of worry; only 5.3 per cent checked 10 or more items.

WILLIAM O. ROBERTSON

cent level in 4 of the 22 categories; at the 1 per cent level in none. Their respective rankings do not differ.

Although variations in, and overlap of, definitions must be considered, the categories designated "digestion," "breathing," and "bowel movements" were noted most commonly as sources of concern. Forty-two per cent of all respondents indicated some degree of worry about "digestion"; only 1 per cent, about "arms and hands." The two categories indicated most frequently as sources of "considerable worry" both related to function or malfunction of the infant's gastrointestinal tract.

A frequency distribution curve of the results from Survey A is shown in Figure 1. The average respondent checked 3.6 items as causes of "some worry" and .5 items of "considerable worry" for a total of 4.1 "worries" per mother.

TABLE 2

COMPARISON OF RESPONSE TO FIRST AND SECOND MAILINGS OF SURVEY A

	<i>N</i>	<i>Mean Number of Worries</i>	<i>SD</i>
1st Mailing	841	4.15	2.87
2nd Mailing	251	3.82	2.98

The mean numbers of concerns elicited in the first and the second mailings of Survey A are compared in Table 2; those from Survey A and Survey B are compared in Table 3. No statistically significant differences were found.

TABLE 3

COMPARISON OF RESPONSE TO SURVEY A AND SURVEY B

	<i>N</i>	<i>Mean Number of Worries</i>	<i>SD</i>
Survey A	1092	4.07	3.06
Survey B	1134	4.03	3.14

The frequency with which a specific "sign" or "symptom" was indicated as a source of concern in Survey A is noted in Table 4. The predominance of concern about functional or physiologic manifestations, as opposed to anatomic ones, is apparent. To be noted is that the fifth most frequently noted cause of worry was "spoiling the baby"; it was the second most frequently noted specific concern among primiparas.

CHILD DEVELOPMENT

TABLE 4

INCIDENCE OF SPECIFIC MATERNAL CONCERNS

(N = 1093 Mothers)

<i>Specific Concern</i>	<i>Percentage of Mothers Indicating Worry</i>	<i>Specific Concern</i>	<i>Percentage of Mothers Indicating Worry</i>
Skin Rash	23.4	Hungry Baby	9.1
Hiccoughs	17.4	Irregular Eating	8.9
Gas	16.7	Vomiting	8.8
Diaper Rash	16.1	Burping	8.6
Spoiling Baby	16.0	Swollen, Protruding Navel ..	8.4
Not Enough Sleep (Baby) ..	12.5	Infrequent, Painful BM	7.9
Too Much Crying	12.1	Choking	7.3
Irregular Sleep	11.9	Accidents While Asleep	7.2
Hard Stools	11.9	Frequent BM	7.0
Spitting Up	11.7	Thumbsucking	6.8
Colic	11.5	Dry Skin	6.8
Restless	11.1	Uneven Breathing	6.7
Loose Stools	10.9	Red Eyes	6.6
Bleeding Navel	9.4	Gags	6.3

To detect the influence that maternal age, uninfluenced by parity, might have in altering the notation of concern, all primiparous respondents were divided into three groups: less than 25 years of age, 25 to 29, and 30 and above. No significant differences were found between the mean numbers of concerns in these three groups.

To determine the influence that level of education might have in altering the mean number of concerns, all primiparous women under 25 years of age were grouped according to level of education; age and parity were thus held constant. No significant difference in concern was found between these groups.

To test for the influence of parity, those mothers 25 to 29 years of age who had completed some, but no more than a high school education, were grouped according to parity as shown in Table 5. As a group, mothers of first children indicated significantly more sources of concerns than did mothers of four or more children. Mothers of two children indicated significantly more concerns than did mothers of three or more children. In support of this finding is the fact that 13 per cent of responding multiparas spontaneously noted under "Comments" words to the effect that their concern had been much greater—both qualitatively and quantitatively—with their firstborn child.

Relatively large and statistically significant decreases in the incidence of certain concerns accompanied increases in parity. Such included "breath-

TABLE 5

RELATION OF PARITY TO MEAN NUMBER OF CONCERNS NOTED BY
MOTHERS 25 TO 29 YEARS OF AGE WITH SOME OR COMPLETE
HIGH SCHOOL EDUCATION

<i>No. of Children</i>	<i>N</i>	<i>Mean</i>	<i>SD</i>
1	51	4.35*	± 2.64
2	120	4.40**	± 2.98
3	121	3.69	± 2.70
4	80	3.20	± 2.97

* Significantly different from mean of parous 4 ($p < .05$).

** Significantly different from mean of parous 3 ($p < .05$) and parous 4 ($p < .005$).

ing," "digestion," and "bowel movements." On the other hand, no significant differences were found with regard to concern about such items as "abnormalities of the navel," "buttocks," or "weight gain." The incidence of worries about "accidents" actually increased significantly with the birth of a second child, but returned to primiparous levels with the presence of four or more children in the family. Of particular interest is the fact that the incidence of concern about colic remained constant (approximately 11 per cent at all levels of parity up through seven); an increase in maternal parity did not alter the frequency of expression of concern about colic.

Variables within the infant such as sex, birth weight, or age in weeks as it ranged from five to ten were not found to be associated with significant differences in the mean number of total concerns indicated. However, several specific concerns did show significant fluctuations in association with sex differences in the infants; these will be discussed in a subsequent paper.

Variation in the geographical location of the respondents, with regard to region of the country or to "urban" or "rural" classification as herein employed, was not found to be associated with significant fluctuations in total number of concerns expressed.

DISCUSSION

Mail Questionnaire Surveys

Prior to interpretation of these results adequate consideration must be given to the certain recognized limitations of the mail questionnaire technique. In this particular instance, certain measures were undertaken to minimize their impact.

1. The written word cannot be assumed to have the same connotation to the reader that the spoken word has for the listener. Neither can it be assumed to have a universal connotation among the members of any one

CHILD DEVELOPMENT

group within the population. Therefore, where possible, potentially confusing or emotionally charged words or phrases were avoided, eliminated, or reduced to a minimum in the questionnaire.

2. The sequential listing of items for consideration is known to influence the testee in completing a questionnaire. In an attempt to minimize this source of error, those sources of concern noted most frequently during preliminary phases of constructing the questionnaire were spread throughout its final form.

3. Clarity of directions for accurate completion of a questionnaire must be included to minimize misinterpretation. Pretesting was, therefore, undertaken to confirm the adequacy of instructions.

4. A precise quantitation by mail questionnaire of the intensity of a specific emotion is obviously impossible. Nevertheless, the respondents were offered the opportunity to note one of two quantifying adjectives, "some" and "considerable." In analyzing the data for correlative relationships, however, the total number of concerns, regardless of indicated intensity, was utilized.

5. While desirable, precise differentiation of concerns based on "objective reality" as contrasted to those based on "subjective fantasy" is also obviously not feasible. Consequently, all concerns, regardless of apparent etiology, were included in the total used for correlative purposes.

6. The absence of a specific category in which to place her source of concern might lead a mother to (a) misplace the source in another category, (b) exclude the specific source altogether, or (c) fail to respond with a completed questionnaire. Consequently, a space for comments was included with the hope that none of these alternatives would be realized.

7. A questionnaire based on leading questions generally elicits a greater total number of answers than does one based on open-ended questions. By calling attention to specific items, it may actually entice the respondent to check an excess number of items. Such must be considered in interpreting the results of questionnaires; they reflect *relative* rather than *absolute* values.

8. Use of a selected group for sampling without an awareness of this by the tester invalidates any generalizations derived from resultant data. The group sampled was recognized as being selected on the basis of their all being mothers of young infants and their all having mailing addresses. Similarly, respondents automatically become still further selected in that they returned the questionnaire. No direct information is, therefore, available regarding the group that failed to respond. However, because of the consistency in incidence of concern between responses to the first and second mailings, further extrapolation from the data may possibly be warranted.

In spite of these inherent limitations, the mail questionnaire survey, by eliminating certain variables that pertain to the personal interview, can still provide meaningful data provided a high percentage of returns is obtained. The percentage of returns in both surveys ($60.5 + 47.5 = 53$ per cent)

is considerably higher than that anticipated in a mail questionnaire survey which is largely nonselected. The primary reason, no doubt, is probably to be found in the high degree of interest possessed by potential respondents in the topic under survey. Admittedly, a still higher percentage of returns would be desirable, but the comparability of replies to first and second mailings suggests that, for the purposes proposed, the technique is satisfactory.

That the second survey, employing an identical questionnaire, duplicated the results of the first so closely evidences the reproducibility of findings. Also, the fact that for both surveys the distribution of returns compared closely with that of control data (except with regard to level of education) would suggest this technique constitutes an economical but unexploited method for gathering certain childrearing data from widespread sources during a brief span of time.

Mothers' Worries

As illustrated in Table 1, the relative frequency with which responding mothers noted various sources of concern is in general agreement with previous more limited observations. Carithers (3), in tabulating the concern among mothers in his practice, as these were expressed during the period of hospital confinement, found that "physiologic" outnumbered "anatomic" areas as sources of concern by a ratio of 4:1. In his study, worry about the infant's gastrointestinal tract and the skin was also most common. Montgomery (7), too, found the gastrointestinal tract to be the central focus of concern. When he contrasted the incidence of "chief complaints" among a group of mothers during the period from 1926 to 1930 with that from another group of mothers during the period of 1946 to 1949, he found that six leading complaints in the first group all related to gastrointestinal function and that five of these six categories continued to predominate as sources of concern in the latter group.

On the other hand, Escalona (4) has found that the baby's crying was the most frequent source of concern. But, as she has pointed out, this is often interpreted by the mother as representing hunger and hence is assumed by her to represent a gastrointestinal aberration. Escalona also notes the frequent finding of other concerns relating more directly to the gastrointestinal tract, such as the infant's vomiting, constipation, etc. She adds that, in this culture, mothers seem better able to relate these particular concerns to their physicians than concerns about the baby's breathing, bathtime, or startle reflex.

Blum (1), working in a well baby clinic that supervised care during the first year of life, also found that "hunger" and "excessive crying" were the most frequently expressed sources of concern. Again, both of these were interpreted by the mothers in her series as relating to gastrointestinal function or malfunction.

CHILD DEVELOPMENT

It would seem pertinent to contemplate the influence that medicine in general, and pediatrics in particular (with its origins so closely intertwined with the development of "safe infant nutrition"), has had or continues to have in possibly encouraging the mother to focus her own anxiety on the infant's gastrointestinal tract. Perhaps a reorientation emphasizing other factors in the infant or in his external environment would be advisable.

Influence of Maternal Factors

Among the maternal factors—age, level of education and parity—only parity was significantly associated with fluctuation in expression of concern. An increase in maternal parity was reflected in a reduction of concerns expressed. Blum has also reported that, on the basis of the total concerns they expressed, multiparous mothers tended to be less anxious or worried than primiparous mothers. Carithers, however, concluded multiparity was not associated with a decrease in total concerns as expressed by subjects in his series. In a differing approach to the influence of parity, Brody (2) separated a group of mothers into four tentative subgroups on the basis of extensive observational and psychological testing criteria. In subgroup A the mothers "were conspicuous for their ability to accommodate to the needs of their infants," while those in subgroup D "were conspicuously active but also erratic in their attentiveness, efficiency and sensitivity. They quite sedulously governed their infants' actions by stimulating, restricting or instructing them, apparently hardly aware of the possible effects of their behavior on the infants' condition . . ." (pp. 265-266). Of interest is the fact that among the seven mothers in subgroup A only one was primiparous, while among those in group D eight of eleven were primiparous, a statistically significant contrast. The present investigation would support her interpretation that with increase in parity, and consequently in experience, there is a decrease in expressed maternal anxiety. How this may relate to currently popular theories on the etiology of colic, the incidence of which remained constant at all levels of parity will be discussed in a subsequent paper.

Among the specific concerns, the observed initial increase in concern about accidents associated with an early increase in parity, and its subsequent return to primiparous levels as parity increased, would seem to reflect experiential learning on the part of the mother. The marked decrease in incidence of concern about bowel movements, sleeping, spoiling, and crying associated with increase in parity would also suggest that learning by experience can be successful in reducing a mother's anxiety. On the other hand, the relative stability of concern about the navel or the buttocks would suggest that, in spite of the mother's learning by experience, specific items continue to represent sources of concern to her. Such a finding makes more understandable the observation that, while the physician's brief explanation about one source of concern may calm an anxious mother, his detailed and repeated

WILLIAM O. ROBERTSON

explanations about another may be totally without effect. It also suggests that, while anticipatory instruction or education about certain aspects of baby care or childrearing might be of considerable value, that about others would prove of little or no benefit.

While cautious interpretation is no doubt called for, other implications may be recognized in the results of this investigation. The absence as well as the presence of associations form the basis for many postulates. For example, the lack of association between fluctuation in the total number of concerns expressed by mothers and variables within their infants suggests that the focus of child health supervision might profitably include more of the mother. It would seem to confirm beliefs that she, rather than he, accounts for much of the contrast in emphasis of the infant's history. In such instances, therapeutic intervention aimed solely at the infant as the source of concern would be doomed to failure.

Recently Ross, in response to an editorial plea to pediatricians for "truly competent counselling in mental hygiene" (6), has emphasized the tenuousness of any claims for a scientific basis in such endeavors, pointing out that the available reference literature "is full of conjecture, speculation and personal impression" and derived from "relatively limited socio-economic groups" (8). When scrutinized, much currently popular childrearing advice—often given under the guise of child health supervision—fares little better, in terms of validity, than does preventive psychiatry. It, too, frequently stems from contradictory proclamations of intuitive dogma unsupported by data. It is suggested that selected use of the mail questionnaire survey—in addition to more widely accepted devices—be employed more extensively in attempts to gather objective data in this field. Such might preclude the obvious "reversal of field" in childrearing advice that has transpired in the past two generations from occurring in the future.

SUMMARY AND CONCLUSIONS

Two mail questionnaire surveys were conducted among nationwide samples of mothers of young infants to determine the incidence of their various concerns. The results were comparable. The mothers were found to be nonselected except, as was anticipated, with regard to their level of education.

Concerns about gastrointestinal dysfunction, skin disorders, and sleep disturbances were most prevalent. Fluctuations in incidence were analyzed for association with such variables as maternal age, education, and parity, infant's sex, and geographical location of the family. Only an increase in maternal parity was associated with a reduction in incidence of concerns.

It is suggested that the technique of mail questionnaire survey offers an effective and economical method for collecting objective data. Wider usage should be considered in approaching problems relating to comprehensive child health supervision.

CHILD DEVELOPMENT

APPENDIX

DUPLICATE OF STUDY QUESTIONNAIRE

Baby's date of birth _____ Today's date _____

Baby's weight at birth _____ Mother's age _____

Baby is a boy ☐ girl ☐ How many *other* children do you have? _____

The following list is based on doctors' reports of the many questions or worries that mothers sometimes have about their new babies. Which of these, if any, have *worried* you about your baby?

	No WORRY (check)	SOME WORRY (check)	CONSIDERABLE WORRY (check)	PLEASE DESCRIBE
STOMACH (too large, small, hard, soft, swollen, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
BREATHING (uneven, hiccoughs, gags, chokes, gasps, grunts, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
HAIR (too much, too little, falling out, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
EYES (puffy, red, crossed, color, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
EARS (shape, size, color, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
NOSE (size, shape, running, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
MOUTH-LIPS (size, shape, color, sore, swallowing, thumb-sucking, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
SLEEPING (too much, not enough, not regular, restless, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
ACCIDENTS (while sleeping, eating, playing, bathing, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
HEAD (size, shape, soft spot, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
WEIGHT (not gaining, too fat, too thin, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
CRYING (too much, too little, strong, weak, turns color, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____

(Study Questionnaire continued on next page)

WILLIAM O. ROBERTSON

DUPLICATE OF STUDY QUESTIONNAIRE (continued from previous page)

NAVEL (swollen, too large, small, bleeding, odor, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
BUTTOCKS (diaper rash, sore, color, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
SKIN (oily, dry, rash, scratches, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
BOWEL MOVEMENTS (odor, color, too often, too loose, hard, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
URINE (odor, color, too often, too little, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
EATING (not enough, too much, not regular, hungry, disagrees, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
DIGESTION (spitting, burping, gas, vomiting, colic, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
LEGS-FEET (too thin, too heavy, not straight, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
ARMS-HANDS (too thin, too heavy, not straight, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____
OTHER (spoiling baby, food preparation, bathing, clothing, diapering, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	_____

Have you chosen a doctor to care for your baby yet? Yes ☐ No ☐

If yes, has the doctor examined your baby in—

(Please check) a. the hospital ☐ b. the office ☐
c. your home ☐ d. baby not examined by doctor yet ☐

Did the doctor mention anything about your baby that would need special care or attention? (Please explain) _____

About yourself, how many years were you in— Grade School _____ yrs.;
High School _____ yrs.; College _____ yrs.; Postgraduate _____ yrs.

Comments _____

CHILD DEVELOPMENT

REFERENCES

1. BLUM, L. H. Some psychological and educational aspects of pediatric practice: A study of well-baby clinics. *Genet. Psychol. Monogr.*, 1950, 41, 3-97.
2. BRODY, S. *Patterns of mothering*. International Universities Press, 1956.
3. CARITHERS, H. A. Psychologic aspects of pediatrics: Mother-pediatrician relationship in the neonatal period. *J. Pediat.*, 1951, 38, 654-660.
4. ESCALONA, S. Problems of infancy and childhood. In M. J. E. Senn (Ed.), *Transactions of third conference*. Josiah Macy, Jr. Foundation, 1949. Pp. 30-95.
5. MAY, C. D. Editorial: Can the new pediatrics be practiced? *Pediatrics*, 1959, 23, 253.
6. MAY, C. D. Editorial: Mental hygiene in the new pediatrics. *Pediatrics*, 1959, 24, 355.
7. MONTGOMERY, Z. D. Problems of infancy and childhood. In M. J. E. Senn (Ed.), *Transactions of third conference*. Josiah Macy, Jr. Foundation, 1949. Pp. 11-29.
8. ROSS, L. S. Letters to the editor. *Pediatrics*, 1960, 25, 174-175.
9. SHAFFER, T. E. *Hospital acquired staphylococcal disease; example of questionnaire survey of discharged infants*. Proc. Natl Conf. Dept. Hlth, Educ. Welf., 1958.
10. SNEGIREFF, L. S., & LOMBARD, O. M. Smoking habits of Massachusetts physicians: Five-year follow-up study (1954-1959). *New England J. Med.*, 1959, 261, 603-604.
11. Annual Report: Vital statistics of the U. S., 1954, Vol. I. Dept. Hlth, Educ. Welf., Public Hlth Service, Natl Office Vital Statistics, 1955.
12. Annual Survey for 1956, Part I, Monthly vital statistics report. Dept. Hlth, Educ. Welf., Public Hlth Service, Natl Office Vital Statistics, 1957.
13. Current population reports. Bureau of Census, Washington, D.C., March, 1957.

POSITIVE REINFORCEMENT AND BEHAVIORAL DEFICITS OF AUTISTIC CHILDREN

C. B. FERSTER*

Indiana University Medical Center

Infantile autism, first described by Kanner (6), is a very severe psychotic disturbance occurring in children as young as 2 years. At least outwardly, this childhood schizophrenia is a model of adult schizophrenia. Speech and control by the social environment are limited or absent; tantrums and atavistic behaviors are frequent and of high intensity; and most activities are of a simple sort, such as smearing, lying about, rubbing a surface, playing with a finger, and so forth. Infantile autism is a relatively rare form of schizophrenia and is not important from an epidemiological point of view. The analysis of the autistic child may be of theoretical use, however, since his psychosis may be a prototype of the adult's; but the causal factors could not be so complicated, because of the briefer environmental history. In this paper, I should like to analyze how the basic variables determining the child's behavior might operate to produce the particular kinds of behavioral deficits seen in the autistic child. To analyze the autistic child's behavioral deficits, I shall proceed from the general principles of behavior, derived from a variety of species, which describe the kinds of factors that alter the frequency of any arbitrary act (3, 10). The general principles of behavior applied to the specific situations presumably present during the child's developmental period will lead to hypotheses as to specific factors in the autistic child's home life which could produce the severe changes in frequency as well as in the form of his behavior. As an example, consider the effect of intermittent reinforcement, many of the properties of which are comparatively well known from animal experiments. To find how intermit-

* Institute of Psychiatric Research, Indiana University Medical Center, 1100 West Michigan Street, Indianapolis 7.

tent reinforcement of the autistic child's behavior might produce deficits, we would first determine, in the general case, what specific orders of magnitude and kinds of schedules produce weakened behavioral repertoires. The factors in the child's home life could be examined to determine estimates of what kind of circumstances could conceivably cause schedules of reinforcement capable of the required attenuation of the child's behavior. The analysis will emphasize the child's performance as it is changed by, and affected in, social and nonsocial environment. As in most problems of human behavior, the major datum is the frequency of occurrence of the child's behavior. Although the account of the autistic child's development and performance is not derived by manipulative experiments, it may still be useful to the extent that all of the terms of the analysis refer to potentially manipulable conditions in the child's environment and directly measurable aspects of his performance. Such an analysis is even more useful if the performances and their effects on the environment were described in the same general terms used in systematic accounts of behavior of experimental psychology.

Some of our knowledge of the autistic child's repertoire must necessarily come from anecdotal accounts of the child's performance through direct observation. Although such data are not so useful as data from controlled experiments, they can be relatively objective if these performances are directly observable and potentially manipulable. A limited amount of experimental knowledge of the dynamics of the autistic child's repertoire is available through a program of experiments in which the autistic child has developed a new repertoire under the control of an experimental environment (5). These experiments help reveal the range and dynamics of the autistic child's current and potential repertoires. In general, the autistic child's behavior will be analyzed by the functional consequences of the child's behavior rather than the specific form. The major attempt will be to determine what specific effects the autistic child's performance has on that environment and how the specific effects maintain the performance.

SPECIFICATION OF THE AUTISTIC CHILD'S PERFORMANCE

We must first describe the current repertoire of the autistic child before we can describe possible environmental conditions that might produce gross behavioral deficits. A topographic description of the individual items of the autistic child's repertoire would not, in general, distinguish it from the repertoires of a large number of functioning and nonhospitalized children, except perhaps in the degree of loss of verbal behavior. The autistic child's behavior becomes unique only when the relative frequency of occurrence of all the performances in the child's repertoire is considered. In general, the usual diagnostic categories do not adequately characterize the children in the terms of a functional analysis of behavior. Hospitalization of a child usually depends upon whether the parent can keep the child in the home,

rather than a functional description of the role of the parental environment in sustaining or weakening the child's performance.

Range of Performances

Although the autistic child may have a narrower range of performances than the normal child, the major difference between them is in the relative frequencies of the various kinds of performances. The autistic child does many things of a simple sort—riding a bicycle, climbing, walking, tugging on someone's sleeve, running, etc. Nevertheless, the autistic child spends large amounts of time sitting or standing quietly. Performances which have only simple and slight effects on the child's environment occur frequently and make up a large percentage of the entire repertoire, for example, chewing on a rubber balloon, rubbing a piece of gum back and forth on the floor, flipping a shoelace, or turning the left hand with the right. Almost all of the characteristic performances of the autistic child may be observed in nonhospitalized children, but the main difference lies in the relative importance of each of these performances in terms of the total repertoire. Conversely, isolated instances of quite "normal" performances may be seen in the autistic child. Again, the relative frequency of the performances defines the autistic child.

Social Control over the Child's Performance

The major performance deficits of the autistic child are in the degree of social control: The kinds of performances which have their major effects through the mediation of other individuals.

The main avenue of social control in a normal repertoire is usually through speech, a kind of performance that is unique because it produces the consequences maintaining it through the mediation of a second person (12). Autistic children almost always have an inadequately developed speech repertoire, varying from mutism to a repertoire of a few words. Even when large numbers of words are emitted, the speech is not normal in the sense that it is not maintained by its effect on a social environment. When normal speech is present, it usually is in the form of a *mand* (12). This is a simple verbal response which is maintained because of its direct reinforcement, e.g., "Candy!" "Let me out." The main variable is usually the level of deprivation of the speaker. It lacks the sensitive interchange between the speaker and listener characteristic of much human verbal behavior, as for example, the *tact* (see below). The reinforcement of the *mand* largely benefits only the speaker. In the case of the autistic child, it frequently affects the listener (parent), who escapes from the aversive stimulus by presenting a reinforcing stimulus relevant to the child's *mand*. At suppertime, the child stands at the door screaming loudly and kicking the door because the ward attendants in the past have taken the child to supper when this situation became aversive enough. Sometimes, the form of the *mand* is nonvocal, although still verbal, as when the *mand* involves tugging at a sleeve, push-

CHILD DEVELOPMENT

ing, or jostling. The dynamic conditions which could distort the form of a mand into forms most aversive to a listener will be described below. In contrast to the mand, the tact (12) is almost completely absent. This form of verbal behavior benefits the listener rather than the speaker and is not usually relevant to the current deprivations of the speaker. This is the form of verbal behavior by which the child describes his environment, as, for example, "This is a chair"; "The mailman is coming." This latter kind of verbal control is generally absent or weak, as with other kinds of verbal behavior except an occasional mand.

Atavisms

Tantrums, self-destructive behavior, and performances generally aversive to an adult audience are relatively frequent in the autistic child's repertoire. Most autistic mands depend on an aversive effect of the listener for their reinforcement. To the extent that social behavior is present at all, its major mode is through the production of stimuli or situations which are aversive enough so that the relevant audience will escape or avoid the aversive stimulus (often with a reinforcer). For example, on the occasion of candy in the immediate vicinity, the child screams, flails about on the floor, perhaps striking his head, until he is given some candy. There is evidence that much of the atavistic performance of the autistic child is operant, that is, controlled by its consequence in the social environment. The operant nature of the autistic child's atavisms is borne out by experiments where a child was locked in an experimental space daily for over a year. There was no social intervention, and the experimental session was usually prolonged if a tantrum was underway. Under these conditions, the frequency of tantrums and atavisms declined continuously in the experimental room until they all but disappeared. Severe tantrums and attempts at self-destruction still occurred when sudden changes in the conditions of the experiment produced a sudden change in the direction of nonreinforcement of the child's performances. Severe changes in the reinforcement contingencies of the experiment produced a much larger reaction in the autistic than in the normal child. Consequently, we learned to change experimental conditions very slowly, so that the frequency of reinforcement remained high at each stage of the experiment. Much of the atavistic behavior of the autistic child is maintained because of its effect on the listener.

Reinforcing Stimuli

The reinforcers maintaining the autistic child's performance are difficult to determine without explicit experimentation. Small changes in the physical environment as, for example, direct stimulation of the mouth, splashing water, smearing a sticky substance on the floor, breaking a toy, or repeated tactile sensations, appear to sustain the largest part of the autistic child's repertoire. Nevertheless, these may be weak reinforcing stimuli which appear to be strong, because the response produces its reinforcement continu-

ously and because alternative modes of responding are also maintained by weak reinforcers. The durability and effectiveness of a reinforcer can usually be determined best by reinforcing the behavior intermittently or by providing a strong alternative which could interfere with the behavior in question. In the controlled experiments with autistic children, most of the consequences we supplied to sustain the children's performance, such as color wheels, moving pictures, music, and so forth, were very weak reinforcers compared with food or candy. Food generally appeared to be an effective reinforcer, and most of the performances associated with going to the dining room and eating are frequently intact. In contrast, the normal children could sustain very large amounts of behavior through the nonfood reinforcements. It is difficult to guess the potential effectiveness of new reinforcers, because the estimate depends upon some performance being maintained by that reinforcer.

In the everyday activities of the autistic children, little behavior was sustained by conditioned or delayed reinforcers. But, in a controlled experimental situation, such activities could be sustained by explicit training. For example, (a) The sound of the candy dispenser preceding the delivery of candy served as a conditioned reinforcer. The fine-grain effects of the schedules of reinforcement show this. The difference in performance produced by two different schedules of reinforcement could have occurred only if the effective reinforcer were the sound of the magazine rather than the delivery of a coin. The actual receipt of the coin or food is much too delayed to produce the differences in performances under the two schedules without the conditioned reinforcer coming instantly after a response. (b) With further training, the delivery of a coin (conditioned reinforcer) sustained the child's performances. The coin, in turn, could be used to operate the food or nonfood devices in the experimental room. (c) Still later, coins sustained the child's performance even though they had to be held for a period of time before they could be cashed in. The child worked until he accumulated five coins, then he deposited them in the reinforcing devices. (d) Even longer delays of reinforcement were arranged by sustaining behavior in the experimental room with a conditioned reinforcer as, for example, a towel or a life jacket which could be used later in the swimming pool or in water play after the experimental session terminated. The experimental development of these performances shows that, even though the usual autistic repertoire is generally deficient in performances sustained by conditioned reinforcement and with delay in reinforcement, the children are potentially capable of developing this kind of control.

Little of the autistic child's behavior is likely to be maintained by generalized reinforcement, that is, reinforcement which is effective in the absence of any specific deprivation. A smile or parental approval are examples. The coins delivered as reinforcements in the experimental room are potentially generalized reinforcers, since they make possible several performances under the control of many different deprivations. However, we

CHILD DEVELOPMENT

do not know whether the coin has actually acquired the properties of a generalized reinforcer.

Stimulus Control of Behavior

It is very difficult to determine the stimulus and perceptual repertoire of autistic children. When a child responds to a complex situation, it is not usually clear what aspect of the situation is controlling the child's behavior. In most cases, it is difficult to determine to what extent these children can respond to speech discriminatively, since the situations are usually complex and many stimuli may provide the basis for the simple performances. Similarly with visual repertoires. Controlled experiments showed unequivocally that behavior can come under the control of simple stimuli when differential effects of the performances were correlated with the different stimuli. When a coin was deposited in a lighted coin slot, it operated the reinforcing device. Coins deposited in unlighted slots were wasted. The children soon stopped putting coins in the unlighted slots. The previously developed stimulus control broke down completely when these stimuli were placed in a more complicated context, however. A new vending machine was installed with eight columns, eight coin lights, and eight coin slots, so the child could choose a preferred kind of candy. The slight increase in complexity disrupted the control by the coin light, and it took several months and many experimental procedures before the stimulus control was reestablished. A better designed procedure, in view of the minimal perceptual repertoire of these children, would have been a gradual program by which variations in the specific dimensions of the coin slot and coin light were changed while the reinforcement contingency was held constant in respect to the essential property.

In summary, the repertoire of the autistic child is an impoverished one. Little is known about the perceptual repertoire, but the available evidence suggests that it is minimal. The absolute amount of activity is low, but this deficit is even more profound if the specific items of activity are evaluated in terms of whether they are maintained by significant effects on a social or even nonsocial environment. Most of the child's performances are of a simple sort, such as rubbing a spot of gum back and forth, softening and twisting a crayon, pacing, or flipping a shoelace. Those performances in the child's repertoire having social effects frequently do so because of their effects on the listener as aversive stimuli. Atavisms and tantrums are frequent.

THE EMERGENCE OF PERFORMANCE DEFICITS DURING THE EARLY DEVELOPMENT OF THE AUTISTIC CHILD

Having characterized the autistic child's repertoire, the next step is to determine the kinds of circumstances in the early life of these children which could bring about the behavioral deficits. The general plan is to

state how the major behavioral processes and classes of variables can drastically reduce the frequency of occurrence of the various behaviors in the repertoire of any organism. Then, the parental environment will be examined to determine circumstances under which the actual contingencies applied by the parental environment to the child's behavior could weaken the child's performance similarly. The datum is the frequency of occurrence of all of the acts in the child's repertoire, and the independent variables are the consequences of these acts on the child's environment, particularly the parental environment. All of the terms in such a functional analysis are actually or potentially directly observable and manipulable. In general, the performances in the child's repertoire will be simultaneously a function of many factors, each contributing to changes in the frequency of the relevant performances. It is important, therefore, to consider relative changes in frequency rather than simple presence or absence of various performances. The datum is the frequency of occurrence of the behavior. In the same vein, singly identifiable factors may be interrelated and functioning simultaneously.

The major paradigm for describing the behavior of an organism is to specify the consequences of the act (reinforcement) which are responsible for its frequency. In this sense, the major cause of an instance of behavior is the immediate effect on the environment (reinforcement). The continued emission of the verbal response "Toast" depends on its effect on the parent in producing toast. Every known behavioral process influencing the frequency of a positively reinforced performance is relevant to the problem of defining conditions under which we may produce a behavioral deficit. Given the variables which maintain it, a performance may be weakened by their absence or by changing the order of magnitude. It is perhaps surprising to discover that large behavioral deficits are plausible without any major appeal to punishment or suppression of behavior by aversive stimuli.

Intermittent Reinforcement and Extinction

Intermittent reinforcement and extinction are the major techniques for removing or weakening behavior in a repertoire. The most fundamental way to eliminate a kind of behavior from an organism's repertoire is to discontinue the effect the behavior has on the environment (extinction). A performance may also be weakened if its maintaining effect on the environment occurs intermittently (intermittent reinforcement).¹ Behaviors occurring because of their effects on the parent are especially likely to be weakened by intermittent reinforcement and extinction, because the parental

¹ The reader may suggest at this point an apparent contradiction with the fact that extinction after intermittent reinforcement is more prolonged than after continuous reinforcement. This aspect of intermittently reinforced behavior's durability is not a general proposition, however, and does not hold for behavior which is still being maintained. Behavior reinforced intermittently will, in general, be emitted less frequently and be more easily weakened by emotional factors, changes in deprivation, punishment, and physiological disturbances than continuously reinforced behavior.

CHILD DEVELOPMENT

reinforcements are a function of other variables and behavioral processes usually not directly under the control of the child. The reinforcement of the verbal response, "Give me the book," may go unreinforced because of many factors which determine the behavior of the listener. He may be preoccupied, listening to someone else, disinclined to reinforce, momentarily inattentive, etc. In contrast, the physical environment reinforces continuously and reliably. Reaching for a book is usually followed by the tactile stimulation from the book. Verbal behavior, particularly, depends entirely for its development and maintenance on reinforcements supplied by an audience (usually a parent). Because of the possibility of prolonged extinction and infrequent, intermittent reinforcement, speech and social behavior are the most vulnerable aspects of the child's repertoire. The young child is particularly vulnerable to the extinction and intermittent reinforcement occurring in social reinforcement because only the parental environment mediates nearly all of the major reinforcers relevant to his repertoire. Large parts of the child's repertoire are reinforced by first affecting a parent who in turn produces the reinforcer for the child. The 2-year-old child who asks for a cookie from a parent and gets no response usually has no alternative audience who will reinforce this vocal behavior. The result will either be the extinction of the child's verbal behavior or the reinforcement of nonvocal verbal forms when the child produces a cookie by a tantrum from which the parent escapes by giving the cookie.

Factors in the Parental Repertoire Affecting the Frequency of Reinforcement of the Child's Performances

To find the conditions under which the child's repertoire will be weakened, therefore, we must look for conditions influencing the parents' behavior, which will alter the parental performances, in turn providing reinforcement of the child's performances. These might be:

1. The general disruption of the parental repertoire. Any severe disruption of the parental repertoire will severely affect the frequency with which the parent reinforces the behavior of the child. Consider, for example, the depressed parent whose general level of behavior is very low. One consequence of this low level of behaving will be a lessened frequency of reacting to the child. Therefore, many items in the child's repertoire will be less frequently reinforced in the depressed than the normal parent. The verbal responses, "May I have some bread" or "I want to go outside," might go unreinforced or be emitted many times without reinforcement. Various kinds of somatic disturbances, such as alcoholic "hangover," drug addiction, severe headache, somatic diseases, etc., could also produce large changes in the over-all reactivity of the parent to a child. To the extent that the child's performances occur because of their effect on the parent, the severely weakened parental repertoire may correspondingly weaken the child's behavior. If the parental extinction of the child's behavior is systematic and periodic, much of a child's behavior could be eliminated.

2. Prepotency of other performances. Whether or not a parent reinforces a child's performance also depends upon the alternative repertoire available to the parent. For example, the parent who is absorbed in various kinds of activities such as housecleaning, a home business, social activities and clubs, active telephoning, and so forth, may at various times allow many usually reinforced performances to go unreinforced. In general, the likelihood of omitting reinforcement would depend upon the strength of the prepotent repertoire. As an example of a prepotent repertoire, the housewife absorbed in a telephone conversation will not be inclined to answer a child or comply with a request. Housecleaning might be another repertoire controlling some parents' behavior so strongly that it is prepotent over behavior in respect to the child. In both cases, the essential result is the nonreinforcement of the child's behavior in competition with the prepotent parental repertoire. Mothers of autistic children often appear to have strong repertoires prepotent over the child. This may be at least a partial reason why mothers of autistic children are so often well-educated, verbal, and at least superficially adequate people.

3. A third factor producing intermittent reinforcement of the child's behavior is related to the first two factors listed above. If the parent finds other reinforcers outside of the home more rewarding than dealing with the child, the child becomes an occasion on which the significant elements of the parental repertoire cannot be reinforced. A parent changing diapers, or otherwise taking care of a child, cannot telephone a friend, be out socializing, be on a job, or doing whatever the autistic mother finds rewarding. The child acquires the properties of a conditioned aversive stimulus because it is an occasion which is incompatible with the parents' normal repertoire. This is of course the major method of aversive control in human behavior—the discontinuation of positive reinforcement. Another basis for establishing the child as a conditioned aversive stimulus to the parent is the emergence of atavisms and a large degree of aversive control of the parent by the child. To the extent that the parent is reinforced by escaping from the child because of his conditioned aversive properties, the frequency of the parental reinforcement of the child's behavior is further reduced.

The development of the atavistic behavior in the child by the parent is necessarily a very gradual program in which the beginning steps involve small magnitudes of behavior such as whining, whimpering, and crying. As the parent adapts to these or becomes indifferent to them because of the prepotence of other kinds of activity, then progressively larger orders of magnitude become reinforced. The large-magnitude tantrum may be approximated or "shaped" by gradual differential reinforcement. The parents of one autistic child, for example, at one period took turns all night standing in the child's room because one step out of the room would immediately produce a severe tantrum in the child. When the child functions as a conditioned aversive stimulus for the parent, the parent is less likely to reinforce the child's behavior positively. This lack of positive reinforcement,

in turn, emphasizes the atavistic responses on the child's part as the major mode of affecting the parent.

The usual limiting factor in preventing excessive development of tantrums is the emergence of self-control on the part of the parent in escaping from the aversive control by the child rather than reinforcing it. Here, again, the repertoire of the parent is relevant. The development of self-control requires a highly developed repertoire which depends for its development on the ultimate aversive consequences of the child's control of the parent. The child's control becomes more aversive to the parent if it interrupts strong repertoires. Specifically, a parent engrossed in a conversation will find a child's interruption more aversive than a parent who is simply resting. If, in fact, there is no strong behavior in the parent, then the child's control is not likely to be aversive, and there is no basis for developing self-control.

All three of the above factors—over-all disturbances in the parental repertoire, prepotent activities, and escape from the child because of his aversiveness—reduce the amount of parental reinforcement of the child's performances. The over-all effect of the nonreinforcement on the repertoire of the child will depend upon the length of time and number of items of the child's repertoire that go unreinforced, as well as the existence of other possible social environments that can alternatively maintain the child's behavior (*see below*).

Differential Reinforcement of Atavistic Forms of Behavior by the Parent

The schedule by which the parent reacts to the child is also relevant to the development of atavistic behavior. Initially, a tantrum may be an unconditioned consequence of parental control as, for example, sudden nonreinforcement or punishment. Eventually, however, the child's tantrums may come to be maintained by their effect on the parental environment, because they present an aversive situation that can be terminated if the parent supplies some reinforcer to the child. The reinforcer presented by the parent to escape from the aversive consequences of the tantrum also increases the subsequent frequency of atavistic responses.

The effect on the parent of the given form and intensity of tantrums will vary from time to time, depending on the conditions maintaining the parents' behavior. This variation in sensitivity of the parent to aversive control by the child results in a variable-ratio schedule of reinforcement of the child's tantrum by the parent—a schedule of reinforcement potentially capable of maximizing the disposition to engage in tantrums. This is the schedule of reinforcement that produces the high frequencies of performances as in gambling (10). The sensitivity of the parent to aversive control by the child will depend on the general condition of the parental repertoire as discussed above. The same factors in the parental repertoire that tend to produce nonreinforcement of the child's behavior—general disruption of the parent or other behaviors prepotent over the child—correspondingly

produce reinforcement of large-order-of-magnitude tantrums. The parent whose total repertoire is severely enough disrupted to interfere with the normal reinforcement of the child's behavior will also react only to tantrums that are of large order of magnitude of aversiveness. A range of sensitivity of the parent to aversive control by the child produces ideal conditions for progressively increasing the intensity or frequency of tantrums. A high sensitivity to aversive control guarantees that some tantrums will be reinforced at least periodically. A low sensitivity differentially reinforces tantrums of large orders of magnitude. At one extreme, the parent may be hypersensitive to the child and, at other times, so depressed that only physical violence will produce a reaction. The schedule by which the parent's behavior terminates the tantrum is a second factor which will increase the range of reactivity of the parent. As more behavior is required of the parent to terminate the tantrum, the parent's inclination to do so will fall. When the parent is less inclined to reinforce a given intensity of tantrum, any variation in tantrum intensity is tantamount to differential reinforcement of extreme forms, if the parent now reacts to the larger-order-of-magnitude tantrum.

How much the parent differentially reinforces tantrums in the child depends, in part, upon the child's other positively reinforced repertoires. When, for example, a child's performance suddenly goes unreinforced, as when a parent may refuse a request, the likelihood and severity of a tantrum will in part depend on the parent's ability to "distract" the child. This, in turn, depends upon whether alternative modes of behavior are in fact available to the child. When conditions are present for the progressive reinforcement of more and more severe tantrums, the process is potentially non-self-limiting. Autocatalysis is likely to occur, particularly if the parent has little disposition to reinforce the general items in the child's repertoire for reasons other than terminating the aversive demands of the child.

Nonsocial Reinforcers

Some of the child's behavior is maintained by his direct effect on the physical environment without the intervention of other individuals. In general, very small effects on the environment will sustain performances with which the parent usually has little reason to interfere. For example, the child plays with his own shoelace, moves his fingers in his own visual field, emits minimal nonverbal, vocal responses, and so forth. Larger effects on the physical environment as, for example, moving objects about the house, speaking to the parent, playing with toys, touching and handling usual household objects, are more likely to enter upon the parental repertoire and so may produce a response whose effect is to discontinue the behavior or interfere with its reinforcement. The punishment aspect of the parental interference with the child's activities will be dealt with separately below. The relative possibility of parental interference and nonreinforcement of the hierarchy of performances may account for the large part of

the autistic child's repertoire, which consists of behaviors having small, limited effects on the physical environment. Occasionally, even behaviors that are maintained by the most simple effects on the environment are extinguished or punished when they occur in the presence of a parent. For example, the father of one autistic child reports that the child reached for a chandelier while he was holding him. The father instantly dropped the child, with a reaction of considerable disapproval because "You should pay attention to me when you're with me." Aside from the secondary effect on the child, the immediate result of the incident is the nonreinforcement of the child reaching for a common physical object.

The existence of "nonverbal" vocal behavior in some autistic children may be related to forms of vocal behavior with which the parent will or will not interfere. Vocal behavior maintained by its effect on a parent (verbal) is susceptible to weakening by parental extinction. A parent interferes less easily with vocal behavior maintained by its direct effect (nonverbal) comparable with making noise by rubbing a stick over a rough surface. Further, such nonverbal vocal responses can emerge readily at any stage of the child's life, unlike verbal behavior, because it does not depend on a generalized reinforcement.

Failure to Develop Conditioned and Generalized Reinforcers

The normal repertoire of the child consists almost entirely of sequences of behavior that are maintained, in a chain or sequence, by conditioned and generalized reinforcers (10). An example of a chain of responses would be the behavior of the child moving a chair across the room and using it to climb to a table top to reach a key which in turn opens a cupboard containing candy. This complicated sequence of behavior is linked together by critical stimuli which have the dual function of sustaining the behavior they follow (conditioned reinforcement) and setting the occasion for the subsequent response. The chair in the above example is an occasion on which climbing onto it will bring the child into a position where reaching for food on the table top will be reinforced by obtaining food. Once this behavior is established, the chair in position in front of the table may now be a reinforcer, and any of the child's behavior which results in moving the chair into position will be reinforced because of the subsequent role of the chair in the later chain of behaviors. A minimal amount of behavior is necessary before a chain of responses can develop. The development of the control by the various stimuli in the chain, both as discriminative stimuli setting the occasion for the reinforcement of behavior and as reinforcers, depends upon a high level of activity, so that the responses will occur and come under the control of the stimuli. This is even more true for the development of the generalized reinforcer. When the child has moved enough objects about the house and achieved a variety of effects on his environment relative to a range of deprivations and reinforcers, simply manipulating the physical environment may become a reinforcer without reference to a

specific level of deprivation. This, of course, is the uniquely human reinforcer that makes possible much of verbal behavior, education in general, and self-control. Again, large amounts of behavior—many chains of behavior with many different kinds of conditioned reinforcers—are a necessary condition for the emergence of a generalized reinforcer. To the extent that the child's repertoire becomes weakened by intermittent reinforcement and extinction, as mentioned above, and punishment and aversive control (see below), the possibility of the development of generalized reinforcers, and hence more complex behavior, becomes less and less likely. Parental "attention" is probably one of the most important generalized reinforcers normally maintaining the child's behavior. Parental attention is an occasion upon which the child's performances may have an important effect on the parent. Inattention is an occasion on which the child's responses are likely to have little effect. Hence, the parents' performances in smiling, saying, "Right," "Good boy," or "Thank you," all come to function as conditioned reinforcers. Their emergence as generalized reinforcers again depends upon the existence of a large behavioral repertoire. A large number of chains of responses will produce important positive effects when the parent smiles or says, "Good boy." Lower frequencies of reinforcement follow for these same activities when the parent is frowning or says, "Bad boy."

Any large reduction in the child's over-all performance will interfere with the initial development of conditioned reinforcers or their continued effectiveness. The control by the environment over the child's behavior depends first upon the emission of the behavior. This follows from the manner in which the environment comes to control the child's performance: the successful execution of an act on one occasion, coupled with the unsuccessful act in its absence. Until a child climbs on chairs, as in the previous example, a chair has little chance of becoming a discriminative stimulus. Without the development of stimulus control, conditioned reinforcers cannot develop. The reinforcing effect of the chair in the above example depends upon its being the occasion on which further performances may be reinforced. In this way, a low general level of behavior may impede the enlargement of the child's repertoire because it does not allow stimulus control and in turn prevents reinforcement of new behavior. A limited development of simple conditioned reinforcers in turn prevents the development of a generalized reinforcer. Parental responses, such as smiling, "Good," or "Right," can have little effect on the child if there is not a history by which many different forms of the child's performance have produced various reinforcers on these occasions. Without parental generalized reinforcement, educational processes and positive parental control are all but impossible. This control is normally carried out by the use of praise and parental attention, coupled with mild forms of threats of discontinuing the reinforcers. Even after a generalized reinforcer has acquired its function, its continued effectiveness depends on the various stimuli continuing to stand in a significant relation to the child's performance. The actual form of the parents' generalized

CHILD DEVELOPMENT

reinforcer is not nearly as important as the parents' subsequent reinforcement practices with the child. The reinforcing effects of the smile derive from the reinforcing practices associated with it. A smile usually functions as a generalized reinforcer in most people because a smiling person is more likely to reinforce. The correlation between smiling and reinforcement is by no means inevitable, however. Some individuals may be more disposed to punish than reinforce when smiling in some situations. In a similar vein, if the child has no behavior in his repertoire that will be more likely to be reinforced on the occasion of a parental smile, it matters little what the parent's reinforcing practices are when smiling as against when frowning.

STIMULUS CONTROL

The specific occasions on which a child's performances have their characteristic effects on the environment will subsequently determine whether the child acts. In the absence of the characteristic circumstances under which the behavior is normally reinforced, the child will be less disposed to act in proportion to the degree of similarity with the original situation. Changing a stimulus to one which has not been correlated with reinforcement is another way of weakening a repertoire. New stimuli also elicit emotional responses and general autonomic effects that may interfere with established performances. Here, simply repeated exposure to the stimuli may produce adaptation to the stimuli and eliminate their emotional effects. Ordinarily, the infants' performances are under the control of a limited range of stimuli, usually one or two parents in a limited part of a specific home environment. The discriminative repertoire broadens as the child grows older and other individuals come to be occasions on which his performances have significant effects. The parental environment of the very young child narrows the control of the child's performance to a limited range of stimuli, largely because the parent mediates almost all of the important events affecting the child. A major factor which brings the child's behavior more narrowly under the control of the parent is the nonreinforcement of much of the child's behavior in the absence of the parent. The close control of the child's behavior by the parent weakens the child's repertoire in the absence of a parent much more when there has been explicit differential reinforcement than when there has been simply a limited reinforcing environment.

Sudden shifts in the child's environment may or may not produce major performance deficits. At one extreme, a sudden shift of the stimuli in the child's controlling environment will have little influence if the child already has been reinforced on the occasion of a wide range of circumstances and individuals. At another extreme, a repertoire can be eliminated almost completely if the child has had a history in which major kinds of performances have gone unreinforced except on the occasion of a single person

in a specific environment. The sudden shifts in the situations and persons controlling the child's behavior may occur under a variety of circumstances, such as a sudden change in a constant companion, death of a parent, or a sudden shift in the physical environment. A sudden shift in the environment of one of the subjects reported in the previously mentioned experiment could conceivably have been the major factor in her autistic development. Many of the activities of the child's mother were prepotent over dealing with the child, and she solved the problem by hiring a teenage baby sitter as a constant companion and nursemaid. After a year, the baby sitter left, suddenly and abruptly, leaving the child with the mother. Within four months, the child began to behave less in general, lost speech, and showed increasing frequency of atavisms. The child's repertoire possibly was under such close control of the baby sitter that the very sudden change to the mother created an environment which in the past had been correlated with nonreinforcement. If the child's behavior were under very narrow control by the baby sitter, because of the nonreinforcement on all other occasions, a sudden shift, as in the loss of the baby sitter, could produce a dramatic deficit in the child's repertoire.

Disruptive Effect of Sudden Stimulus Changes and the Amount, Durability, and Range of Behavior

A novel reinforcing environment will not sustain a child's performance unless the repertoire contains behavior of a sufficient range and durability. The new environment weakens the performance because it nearly always requires slightly different forms of behavior. For example, a new person entering a child's home is not so likely to respond successfully to the incompletely developed verbal behavior of a child as the parent. The possibility of the child's affecting the stranger will depend upon his having verbal responses different from those usually reinforced by the parent and, also, durable verbal behavior that will continue to be emitted under the intermittent reinforcement that is likely to occur. If the child's repertoire is durable and extensive enough so that the verbal response may be repeated several times and supplemented by auxiliary behavior, the child has a greater chance of affecting the new person or of being shaped by him. Similarly with other kinds of social behavior. The wider the range of behavior and the greater the disposition to emit it the more likely that the child's performance will be within the range of responses potentially reinforcing by the new environment.

For a stimulus to acquire control over behavior, the child must first emit behavior in the presence of the stimulus. Consider, for example, the performance of a child at a children's party at which there are lots of toys and games, such as bicycles, swings, and so forth. The likelihood of the child's behavior coming under the control of any of the other children as reinforcers is minimal if the new environment suppresses or makes the

CHILD DEVELOPMENT

child's entire repertoire unavailable because it is a novel stimulus and is an occasion on which the child's behavior has never been reinforced. If the behavior of playing with a swing or riding a tricycle is sufficiently strong that it may be emitted even under the adverse conditions of the very strange party environment, then the simple emission of the previously developed behavior provides a situation under which other children at the party may potentially reinforce or otherwise affect the child's repertoire. Simply the acts of eating cake, candy, or ice cream, or picking up a toy put some of the child's behavior under the control of the new environment. Each new performance which can potentially occur at the party provides a basis for the child's reinforcing some behavior of other children at the party or of his coming under the control of the other children's reinforcers. On the other hand, a sudden exposure to a new environment with a weak and narrow repertoire may produce a severe behavioral deficit. In any case, the child will be much less disposed to go to the party if he had behaved unsuccessfully in the new environment. This lower disposition to attend and engage in the party would in turn make it less likely that the child will emit behavior that would be reinforced in the party environment.

Adaptation

The emotional and elicited autonomic effects of novel environments may also interfere with a child's performances. Adaptation to new environments occurs with gradual exposure. A sudden exposure to a new environment will produce gross emotional and autonomic responses which will in turn interfere with, or even completely suppress, the emission of possible operant behavior potentially reinforcing by the new environment. The rate at which the child is exposed to the new environments will determine the magnitude of disturbance. Exposure to a new environment and adaptation of the emotional responses do not necessarily create the potential basis for responding, however. A repertoire that will make contact with the new environment is also necessary.

Amount of Prior Nonreinforcement

The more closely controlled the child's performances are by specific stimuli, the more likely a sudden shift in the environment will produce a cessation of responding. For example, the child receiving minimal care from a parent probably will be less affected by a sudden shift in environment than a child closely affected and controlled by parental response. It is paradoxical that the parent who responds sensitively to the child's performance may be potentially weakening it more than the parent who exerts little control over the child. It is the alternate reinforcement and nonreinforcement that place the child's behavior narrowly under the control of very specific stimuli so that it is much more vulnerable to sudden changes. The range of stimuli in whose presence the child's behavior goes unreinforced will determine the narrowness of the stimulus control.

CUMULATIVE EFFECTS OF A BEHAVIORAL DEFICIT

The continuous development of more and more complex forms of a child's behavior is normally achieved because the parents and community approximate the required performances. At each stage of the child's development, the community reinforces the child's current repertoire even though it is more disposed to react to small increments in the child's performance in the direction of the required complex performances. Should any of the above processes produce a deficit in performance or an arrest in the development of the child's performance, further development of a repertoire would depend upon the community's relaxing its requirements and reinforcing performances in an older child that it normally accepts only from a younger one. Ordinarily, the reinforcing practices of the community are based on the chronological age and physical development of the child.

Only between the ages of $1\frac{1}{2}$ to 4 years does the parent have sufficient control of the child to weaken his performance to the degree seen in infantile autism. This is a critical period in the child's development during which his behavior is especially susceptible to extinction, because the traditional social pattern in the usual family restricts the child's experience to one or two parents. Before the age of $1\frac{1}{2}$, the child has few performances with which the parent will interfere or that have important effects on the parent. Much of the infant's behavior is maintained by simple and direct effects on its environment. As the child approaches 2 years, the rapid development of a behavioral repertoire, particularly social and verbal behavior, makes possible extinction and other forms of weakening. The effectiveness of the parental environment in weakening the child's repertoire depends upon the availability of concurrent audiences for the child's behavior. In general, the 2-year-old child is limited to the home and comes into increasing contact with other environments as he grows older, perhaps reaching a maximum at school age. The presence of an older sibling might appear to preempt the possibility of a sufficient degree of isolation to account for an aversive behavioral deficit. A sibling could provide an alternative to the parent as a reinforcing environment. The behavioral or functional influence of a sibling would depend on the amount and nature of interaction between the children. For example, an older child might possibly completely avoid the younger one or tend to have the same patterns of reaction as the parent. In many cases, the older sibling has playmates outside the home to the complete exclusion of the younger child. The older sibling, in many circumstances, punishes as well as extinguishes the younger child for any attempted participation in his play. There are very few facts as to the exact nature of the interactions in most cases.

The parent as the sole maintainer of the child's behavior is perhaps even more likely when the child is raised in a rural or isolated community, and perhaps with one of the parents largely absent. The above analysis suggests that a survey of severely autistic children would, in general, show them

CHILD DEVELOPMENT

to be first-born children; or, if other siblings were available, they would have provided little interaction with the child. It also suggests that the child would be raised in a house physically or socially isolated from other families or children such that there were no alternative social environments that could provide reinforcement for the child's behavior. When the child was exposed to both parents, it would be expected that both parents were consistent in their nonreinforcement of the child's performances.

AVERSIVE CONTROL AND PUNISHMENT

It has been possible to describe conditions which might produce major behavioral deficits without dealing with punishment or aversive control. A similar account might present a functional analysis of how performance deficits might occur as a result of aversive control. Many writers have already described some of these factors by extending general principles of aversive control to human behavior (7, 8, 11). For the purposes of the analysis presented in this paper, I should like to restrict the discussion of aversive control to its relation to positive reinforcement. Much of human aversive control is carried out by discontinuing or withdrawing reinforcement (3, 10). For example, a frown or criticism may function as an aversive stimulus because these are occasions on which reinforcements are less likely to occur. Even when corporal punishment is given, it is not clear as to whether the resulting effect on the child's behavior is due to a slap or to the lower inclination of a punishing parent to reinforce. Most parents who spank a child will be indisposed to act favorably toward the child for some period of time subsequently. As a result, one major by-product of frequent punishment may be a larger order of interference with the child's normal repertoire along the lines of the positive reinforcement deficits described above.

The obvious effectiveness of punishment in some kinds of human control appears to contradict experimental findings with animals which show punishment to have only a temporary effect on behavior (1, 2, 9). The role of positive reinforcement factors helps resolve the dilemma. The effectiveness of punishment depends on how strongly the punished behavior is maintained by positive reinforcement. The apparent effectiveness of punishment in the control of children may occur when weak repertoires are punished or when the punishment indirectly produces extinction. Most animal experiments using electric shock as an aversive stimulus have used strongly maintained positively reinforced operant behavior as the base-line performance to be punished. The aversive control might be more effective when the performances to be punished are less strongly maintained.

CONCLUSION

As might be expected from the relatively low frequency of infantile autism, the combination of circumstances hypothesized above would occur only rarely. The above hypothesis provides a framework for investigating

the circumstances surrounding the development of the autistic child. All of the variables that might weaken the behavior of a child are directly or potentially observable. The data required are the actual parental and child performances and their specific effects on each other, rather than global statements such as dependency, hostility, or socialization. Not all of the factors responsible for a child's performance may be present currently. Using retrospective accounts, however, makes it difficult to determine the actual correspondence between the verbal statements of the parent and their actual practices in raising the child. The alternatives are, first, an objective assessment of the child's repertoire in a wide enough range of environments so as to allow an assessment of the nature of the environmental control of the child's current behavior; and, second, actual home observations of the specific social consequences of the child's performances early in the development of the disease.

The same kind of functional analysis can be made for the performance of the adult psychotic although the specific deficits observed in autistic children and their manner of occurrence may not be relevant. In particular, the analysis of the adult's behavior would be more concerned with the factors which weaken behavior already in the repertoire rather than the development of new repertoires as with the analysis of the autistic child's behavior. Maintaining already-established behavior is more at issue in the adult than the initial development of a performance as in the case of the child (3).

REFERENCES

1. AZRIN, N. H. Punishment and recovery during fixed-ratio performance. *J. exp. anal. Behav.*, 1959, 2, 301-305.
2. ESTES, W. K. An experimental study of punishment. *Psychol. Monogr.*, 1944, 57, No. 3 (Whole No. 623).
3. FERSTER, C. B. Reinforcement and punishment in the control of human behavior by social agencies. *Psychiat. Res. Repts.* 10, 1958, 12, 101-118.
4. FERSTER, C. B., & DEMYER, M. K. A method for the experimental analysis of the behavior of autistic children. *Amer. J. Orthopsychiat.*, in press.
5. FERSTER, C. B., & DEMYER, M. K. The development of performances in autistic children in an automatically controlled environment. *J. chronic Dis.*, 1961, 13, 312-345.
6. KANNER, L. Early infantile autism. *J. Pediat.*, 1944, 25, 211-217.
7. MILLER, N. E., & DOLLARD, S. *Personality and psychotherapy*. McGraw-Hill, 1950.
8. ROTTER, JULIAN. *Social learning and clinical psychology*. Prentice-Hall, 1954.

CHILD DEVELOPMENT

9. SKINNER, B. F. *The behavior of organisms*. Appleton-Century-Crofts, 1938.
10. SKINNER, B. F. *Science and human behavior*. Macmillan, 1953.
11. SKINNER, B. F. Some contributions of an experimental analysis of behavior to psychology as a whole. *Amer. Psychologist*, 1953, 8, 69-78.
12. SKINNER, B. F. *Verbal behavior*. Appleton-Century-Crofts, 1957.

COMPARISON OF DATA OBTAINED FROM MOTHERS AND FATHERS ON CHILDREARING PRACTICES AND THEIR RELATION TO CHILD AGGRESSION¹

LEONARD D. ERON,* THOMAS J BANTA,²
LEOPOLD O. WALDER, and JEROME H. LAULICHT³

Rip Van Winkle Foundation

In reviewing the literature having to do with behavior of children as it is related to behavior and attitudes of parents, one is struck by two methodological biases which do not permit a clear-cut evaluation of the substantive findings of a majority of these studies. The first has to do with the disregard of the father both as an important socializing influence and as an informant. A search of the literature between 1929 and 1956 revealed 160 publications dealing with mother-child relationships but only 11 with father-child relationships (10). Since that time there have been a few more studies which have considered the importance of father-centered variables, e.g., Bronfenbrenner (2), Miller and Swanson (9), Sears, Maccoby, and Levin (11), Sears, Pintler, and Sears (12), but information about fathers in these studies usually has been obtained second-hand, either from the mother or from the child himself. This introduces the second methodological bias which is the contamination that is likely when assessments of both parental and child behavior (presumed antecedent and consequent) are made by the same individual. A study in identification by Levin and Sears (8), as well as the earlier Sears' studies (12, 13), are not subject to this error, although they are to the first. When both the predictor and criterion are obtained from the same subject, then response set cannot be eliminated from the possible interpretations of the results. In addition to contributing important information which has usually been lacking in childrearing studies, the use of both mother and father as informants could serve as one check on this source of error. The optimal procedure, of course, is to have completely

* Mental Health Research Center, Rip Van Winkle Foundation, 886 Columbia Street, Hudson, New York.

¹ This is a revised version of a paper delivered at the Eastern Psychological Association meetings, New York City, April 16, 1960. The research has been supported by grant M-1726 from the United States Public Health Service. Thanks for cooperation and support are due to W. Richards Bonneau, Principal, Ockawamick Central School; Robert Cullen, Principal, Claverack Union Free School; George Pike, Elementary Supervisor, Hudson Public Schools; Jack Roosa, Principal, Greenport School; and the following third-grade teachers: Mrs. Kathelyn Caswell, Mrs. Grace Clapp, Mrs. Rose Healy, Mrs. Catherine McNamara, Mrs. Patricia McNamara, and Mrs. Winifred Rockefeller. We should also like to thank all the children and their parents, who of course must remain anonymous, for their excellent cooperation.

² Now at University of Wisconsin.

³ Now at Berkshire Industrial Farm, Canaan, New York.

Child Development, 1961, 32, 457-472.

CHILD DEVELOPMENT

independent observations of the child and individual measures obtained from both parents. A few recent studies have approximated this goal. Davidson *et al.* (5) administered anxiety questionnaires to a group of children; they then interviewed the mothers *vis a vis*, but merely left a check list of personality characteristics of the child to be completed by the father. Both parents were interviewed by Peterson *et al.* (10), but they did not have independent observations of the child's behavior; although they partially overcame this type of contamination by comparing the attitudes of one parent with judgments of the child's behaviors based on reports of the other parent. Despite these limitations on clear-cut interpretation, both of these studies showed differential results when predicting to child behavior from either mother or father interviews. In the study here reported it has been possible to interview fathers and mothers separately and to obtain independent measures of the children from other informants. It is with the relations among items from these three sources of information that this paper is concerned.

SUBJECTS AND PROCEDURES

A group of 158 children, the total population of six third-grade classes from divergent socioeconomic areas in a semirural county, served as a pool of subjects. Sixty children were drawn from this group on the basis of their scores on an aggression sociometric (described below) administered in the classroom. Thus, the subjects included the 10 most aggressive boys in this sample, the 10 least aggressive boys, 10 boys in the middle range, and three similar groups of girls. Both parents (where available) of each of these children were contacted for personal, individual interviews. The families of only three of these preselected children refused to participate, and they were replaced by subjects with very comparable scores. It was thus possible to interview all of the mothers and 50 of the 58 available fathers. The interviews were conducted by six experienced interviewers⁴ either in the subject's home or the Foundation offices. At the time of the interview, the interviewers had no knowledge of the aggression classification of the subject.

MEASURES

Aggression Index

This is a Guess-Who technique in which every child in a class rates everybody else on a selected series of 22 specific aggression and four aggression anxiety items. These 26 items were what remained after a process of "sifting and winnowing" from an original pool of 1000 short behavioral descriptions which had originally been collected from the literature and by interviewing experts:

⁴Thanks are due to Dr. Marjorie Collins and Mrs. Madeline Eron, psychologists, and Miss Mary Lawrence, psychiatric social worker, who assisted three of the authors (LE, JL, and LW) in obtaining these interviews.

THE AGGRESSION INDEX

(The labels in parentheses refer to type of aggression.)

Cluster 1

1. Who is a pest? (Indirect)
2. Who does not obey the teacher? (Unclassified)
3. Who takes the teacher's things without permission? (Acquisitive)
4. Who is always getting into trouble? (Unclassified)
5. Who tattles to the teacher? (Indirect)
6. Who is rude to the teacher? (Verbal)

Cluster 2

7. Who starts a fight over nothing? (Physical)
8. Who says mean things? (Verbal)
9. Who makes it hard for children to get things done? (Indirect)
10. Who pushes or shoves children? (Physical)
11. Who does things that bother others? (Indirect)

Cluster 3

12. Who forgets to return borrowed things? (Acquisitive)
13. Who often says "Give me that"? (Acquisitive)
14. Who makes marks on the desk? (Physical, against property)
15. Who takes other children's things without asking? (Acquisitive)

Individual Aggression Items

16. Who will always fight back if someone else hits them first? (Physical)
17. Who gives dirty looks or sticks out their tongue at other children? (Verbal)
18. Who complains to the teacher when she tells them what to do? (Verbal)
19. Who grabs things from other children? (Acquisitive)
20. Who uses bad words when another child bothers them? (Verbal)
21. Who gets very, very mad at times? (Unclassified)
22. Who makes up stories and lies to get other children into trouble? (Indirect)

Anxiety about Aggression Items

23. Who is always polite? (Unclassified)
24. Who will never fight even when picked on? (Physical)
25. Who will never argue even when they are right? (Verbal)
26. Who says, "Excuse me," even when they have not done anything bad? (Unclassified)

Rejection by Peers

27. Who are the children that you wish were not in your class at all?

Popularity with Peers

28. Who are the children that you would like to have for your best friends?
29. Who do you know best of all?
30. Who would you like to sit next to in class?
31. Who would you like to play with?

They represented the final distillation of a long series of preliminary studies including armchair judgments by experts as to type of aggression involved, pilot studies with groups of 8-year-olds, and an extended tryout of two alternate forms on 974 third graders, with subsequent cluster and item analyses. The purpose was to find items on which the children, as judges, could agree with each other as to who of their classmates did what. Also included in this

CHILD DEVELOPMENT

sociometric were one rejection and four popularity items. The procedure involved giving each child a booklet containing one page for each question. The names of all children in the class were listed on each page, and subjects were instructed to cross out the names of those children who acted in the way described by the particular question for that page which was read aloud to them by the examiner. The score for each child was based on the number of judges choosing him as someone who fitted a particular behavioral description. Thus, if 10 members of a class crossed out Johnny Jones' name as someone who said mean things, Johnny's raw score was 10 for that item. The raw scores were converted into percentages to correct for differences in class size.

The reliability of this instrument, based on interjudge agreement, had been shown to be excellent, with minimum shrinkage on cross-validation. A cluster analysis isolated three sets of items, one concerned with aggression against teacher, one with aggression against peer, and one having to do with a particular type of aggression which was called acquisitive (*see* Table 1). Also included in this measure which was derived from the two alternate forms were a number of individual items, each with good reliability (at least .70). These had been retained in order to tap different types of aggression, different objects, and different provocation levels.

TABLE 1
CLUSTER RELIABILITIES

	<i>Cluster</i>	<i>Median Reliability*</i>	<i>Cross Validation</i>
Teacher as object	1	.92	.86
Peer as object	2	.87	.77
Acquisitive	3	.85	.80

* The median reliability is based on 12 different groups (six classes, each divided by sex into two groups). Typically, the mean is higher than the median.

There was no relation between score on the aggression measure and amount of interpersonal activity the child engages in as rated by the teacher (r ranged from +.22 to -.38 in six classes with median r of .00) nor with the position of a child's name on the list (r ranged from +.19 to -.15 with median r of .00). As for the distribution of the aggression scores among the subjects, it was found that very few children were highly aggressive; however, the vast majority were moderately aggressive and girls had lower scores than boys. The more popular children were less aggressive, but even the most popular children were rated as aggressive in some ways. On the cluster having to do with aggression towards the teacher, identically shaped distributions of scores (J curves) were obtained for boys and girls. This type of distribution would be expected in most kinds of nonconforming behavior, of which aggression towards teachers is certainly an example.

In addition, Cureton's r_{rb} (4) of .84, uncorrected for attenuation, was obtained between teacher's and peers' rating on this cluster. It is reasonable to assume that the high reliability coefficients themselves indicated validity because the raters were probably basing their judgments on observations of the same specific behaviors (3, 6). The shape of the distribution curves and its consistency from class to class and between boys and girls were further evidence of validity, as were the predicted relations into which the aggression and aggression anxiety items entered with the other classroom measures: popularity-rejection, IQ, and teacher's ratings of aggression (see Table 2). There was also some clinical evidence of validity since a few of the subjects had been referred to the Clinic Guidance Service prior or subsequent to the administration, and in each case, the children who were shown to be overly aggressive scored high on the sociometric.

TABLE 2
INTERCORRELATIONS AMONG CLASSROOM MEASURES*

<i>Other Measure</i>	<i>Aggression Measure</i>	<i>Aggression Anxiety</i>
Rejection73	-.46
Popularity	-.38	.48
Teacher rating on all items63	-.34
IQ	-.17	.20

* A correlation of .23 is significant at the .01 level of confidence.

Parent Interview

This is an objective, almost entirely precoded interview which takes about 1½ hours to administer. Its purpose is to get at the sociocultural and psychological antecedents of aggression as they are mediated by child-parent interaction. Included in this interview, which was pretested on 200 mothers and fathers, are a number of variables suggested as important by both general behavior theory and clinical hunches. It was felt that a profitable way to look at the differential distribution of behaviors which exist in varying degree in different segments of society was through the learning environments provided by these subcultures. Thus, the pertinent variables had to do with rewards and punishments for aggressive behavior, standards of aggressive behavior, role models, identification, and presumed frustrating antecedents to such behavior. These variables are listed below, with definitions and examples of items to tap them:

PARENT INTERVIEW VARIABLES

1. *Approval of Aggression*—Evaluative standards for aggressive behavior stated in terms of approval or disapproval of specific items of aggressive behavior which appear in the school aggression measure. Example: Suppose

CHILD DEVELOPMENT

NAME said mean things to another child. Would you: 0. strongly disapprove, 1. mildly disapprove, 2. not care (and don't know), 3. mildly approve, 4. strongly approve.

2. *Confessing by Child*—Extent to which a child behaves as if he were monitoring his own behavior in a way he thinks a socializing agent would. These items are closed-end versions of two questions by Sears, Maccoby, and Levin (11). Example: When NAME has done something naughty and you haven't seen him do it, does he come and tell you about it without your having to ask him? 4. all the time, 3. most of the time, 2. some of the time, 1. almost never, 0. never.

3. *Dependence Avoidance of Child*—Inability or unwillingness of the child to accept help or rely on others. Example: Does NAME seem embarrassed when you take his part? 0. no, 1. sometimes (and don't know), 2. yes.

4. *Ethnicity* (generational level)—Number of generations in which parents' forebears have lived in the United States.

5, 6. *Father's and Mother's Aggression*—Tendency of the parent to display aggressive behavior in situations which often elicit aggressive responses. Each parent rates himself and is rated by the other parent on each of these situations. Example: Suppose you are driving a new car and get into an accident which is clearly the other driver's fault. Would you show your anger if: 1. he says, "What's the matter, can't you drive?" 2. he apologizes? 3. he says, "I'm not going to say anything until I see a lawyer?" 4. he laughs it off?

7. *Home Aggression of Child*—Frequency of acts whose goal response is injury to another object. Example: How often does NAME say mean things to another child? 0. never, 1. rarely, 2. occasionally, 3. pretty often, 4. frequently, 5. daily. These alternatives were specifically defined in terms of frequency in time and typed on cards which were handed to the respondent.

8. *Lack of Social Participation*—Degree of participation of respondent in formal and informal social relationships. Example: About how many times in the past year have you attended meetings or affairs of any local organizations, societies or clubs? 0. 13+, 1. 7-12, 2. 4-6, 3. 1-3, 4. none.

9. *Nonrecognition of Child's Needs*—An aspect of nonnurturant behavior. (For other aspect of nonnurturant behavior see *Punishment for Nurturance Signals* below.) Example: Do you usually have time so that NAME can talk to you about things that interest him? 0. yes, 1. no (and don't know).

10. *Parental Aspirations for Child*—Level of education parent hopes child will attain. Example: How much education do you expect NAME to get? 1. high school + specialized training or college, 2. high school graduate or less.

11. *Parental Aspirations for Self*—Example: When you left school what particular kind of occupation or life work was it your ambition to reach some day? Aspirations for: 1. professional status, 2. minor profession, small business or farm owner, 3. skilled worker trades, 4. semiskilled or unskilled occupations.

12. *Parental Disharmony*—The extent of disharmony in the home as measured by disagreement about various specific matters of importance in a family; items dealing with arguments between husband and wife, and presumptive evidence such as separation, divorce, amount of time spent together, etc. Example: Are you satisfied with how your SPOUSE handles money? 0. yes, 1. sometimes (and don't know), 2. no.

13. *Parental Rejection*—The number of changes in the child's behavior (aggression excluded) and characteristics desired by the socializing agent. The parent is considered to be accepting when he indicates that his needs are

ERON, BANTA, WALDER, and LAULICHT

satisfied by the child: "I like you the way you are." Example: Do you think NAME wastes too much time? 2. yes, 1. sometimes (and don't know), 0. no.

14. *Parental Restrictiveness*—Extent to which the child defines behaviors which are proper for him to perform rather than the agent defining proper behaviors for the child. Restrictiveness refers to the amount of control exercised by the agent over the child. Example: Do you make NAME finish up everything he is served as mealtime? 2. yes, 1. sometimes (and don't know), 0. no.

15. *Punishment for Aggression*—Rewards and punishments of various intensities administered by socializing agents contingent upon the child's aggressive behavior. Example: What do you usually do when NAME is rude to you? (Verbatim response and probes recorded, subsequently rated by three judges on a scale from 1 to 7.) 1. rewarding aggression, 2. don't do anything, 3-7. mild to severe punishment for aggression.

16. *Punishment for Dependency*—Rewards and punishments of various intensities administered by socializing agent when child asks for help. Example: What do you usually do when NAME asks for help? (Each response was rated by three judges on a scale from 1 to 4): 1. giving help, to 4. punishing the child.

17. *Punishment for Nurturance Signals*—Rewards and punishment of various intensities administered to child by socializing agent in situations which might tend to lead to nurturant behavior on part of agent. Example: What do you usually do when NAME is afraid? (Each response was rated by three judges on a scale from 1 to 4): 1. giving nurturance, to 4. punishing the child.

18. *Residential Mobility*—A measure of the number of times the child changed residence and thereby had to change schools and/or find new friends. Example: How often has moving meant that NAME had to find new friends? 0. none, 1. once, 2. 2-3 times, 3. 4+.

19. *Rural Background*—Population size of geographic area in which respondent was born and grew up. Example: Where did the family live when you were born—on a farm, or in a village, town, small city, medium-sized city, or big city? 1. big city (500,000+), 2. medium-sized city (100,000-500,000), 3. small city (10,000-100,000), 4. town (1,000-10,000), 5. village (under 1,000), 6. farm.

20, 21. *Shame, at home and out of home*—Tendency to punish in public assessed by items involving different kinds of punishments and different publics. Example: Suppose NAME was naughty and you felt he deserved a scolding. Would you do it when: (1) your SPOUSE and other children could hear it? (2) one of NAME's friends could hear? (3) one of your close friends or relatives could hear? (4) a neighbor or acquaintance could hear? (5) you were in public and someone else might hear? 0. no, 1. sometimes (and don't know), 2. yes. The first item constitutes a separate measure, punishing the child in front of his family which defines home shame.

22. *Social Isolation of Child*—Frequency and type of contacts with peers outside of school. Example: About how many children of NAME's age live in the neighborhood? Would you say about one or two, three to five, or more than five? 3. none, 2. 1-2, 1. 3-5, 0. 5+.

No definition was accepted for these variables until items could be written which were judged by a number of experts to fit the definition. Emphasis was placed on the necessity for distinctive definitions and measures of each

CHILD DEVELOPMENT

concept in order that hypotheses about how they relate to overt aggression in school could be tested both singly and in combination. The median number of items for each variable was 10, with a range from 1 to 20.

Before the relations of mother and father interview variables to each other and to the independently obtained criterion measure were tested, the homogeneity of the variables was increased by eliminating all items which correlated less than .30 with the total scale score of which it was a part. These item-total correlations were done separately for mothers and fathers since it was felt that there would be some questions more appropriately asked of fathers and others of mothers. For 14 of the 22 variables under consideration, overlap of questions between mother and father was complete. Of the remaining eight, subsequent analyses on six of them showed there was little difference in results whether or not items completely overlapped (1). It turns out that it makes little difference in this report whether scores are calculated on the basis of the refined scales with incomplete overlap between mother and father questions or on scales with identical questions for both groups.

RESULTS AND DISCUSSION

Once homogeneous sets of items had been derived, a number of correlational analyses were done. First, scores on each variable from mother's interview were correlated with scores on similar variables in the father's interview. As is shown in Table 3, of 22 correlations, only 10 were significantly better than zero ($p < .05$). It is obvious that mothers' and fathers' responses cannot be substituted for one another, at least in this childrearing questionnaire.

An examination of Table 3 reveals that the higher correlations generally are for those variables which have to do with the respondent himself, his

TABLE 3
CORRELATIONS BETWEEN MOTHERS AND FATHERS ON
INTERVIEW VARIABLES*

<i>Variable</i>	<i>r</i>	<i>Variable</i>	<i>r</i>
Residential Mobility91	Confessing by Child25
Parental Rejection64	Ethnicity (Generational Level)21
Lack of Social Participation63	Nonrecognition of Child's Needs ..	.20
Parental Aspirations for Child52	Shame Out of Home16
Parental Disharmony46	Mother's Aggression15
Social Isolation of Child42	Punishment for Nurturance Signals	.13
Punishment for Aggression42	Approval of Home Aggression09
Parental Restrictiveness38	Parental Aspirations for Self07
Father's Aggression35	Shame at Home	-.06
Rural Background32	Child's Dependence Avoidance ...	-.05
Home Aggression of Child25	Punishment for Dependency	-.04

* $r_{.05} = .28$.

own attitudes and feelings; the lower correlations, for the most part, are for those variables having to do with the child's behavior or the parent's interaction with the child. This reveals something about marital choice perhaps (likes attract), but it also shows that mother and father cannot agree in reporting their perceptions of their own child's behavior. However, on the one really objective item that both parents rated, residential mobility, the correlation was .91. Thus, it is not that the parents cannot render reliable information, but perhaps mother and father each observe and react to children differently, and the observations and reactions of each must be taken into account to get a complete picture of parental socialization influences on child behavior.

The same conclusion can be drawn from the next analysis which was the intercorrelation of each interview variable with the other, as well as with the independently obtained classroom measures, done separately for mothers and fathers. For any one variable, all correlations significant at or beyond the .05 level for either mother or father are reported. These appear in Table 4. Although the correlation between mothers' and fathers' scores on any one variable may be better than zero, their scores on that particular variable do not relate to those on other variables in the same way. Thus, for example, the correlation between mother's and father's rating of rejection is .64; however, while mother's rejection score relates significantly to all the classroom measures, father's rejection relates to only one of them. On the other hand, father's rejection relates to a number of other variables in his interview while mother's rejection relates to none of these.

The opposite is true for ethnicity or generational level, a classification of parents on how long ago their ancestors came to this country (on which

TABLE 4
SOME CORRELATIONS AMONG SCALES

Variable	Mother N = 59 $r_{.05} = .26$	Father N = 50 $r_{.05} = .28$
A. Parental Rejection (r mother-father = .64)		
School Aggression	.40	.31
Aggression Anxiety	-.37	-.17
Popularity with Peers	-.29	-.12
Peer Rejection	.35	.20
Home Aggression of Child	.11	* .49
Confessing by Child	-.07	** -.58
Parental Disharmony	.10	* .42
Parent's Aggression	-.03	* .32
Rural Background	-.12	-.37
Punishment for Dependency	.07	.36
Punishment for Aggression	.20	.44

(continued on next page)

CHILD DEVELOPMENT

TABLE 4 (continued)
SOME CORRELATIONS AMONG SCALES

Variable	Mother N=59 $r_{.05} = .26$	Father N=50 $r_{.05} = .28$
B. Ethnicity (generational level) (r mother-father = .21)		
School Aggression	.00	.26
Popularity with Peers	.08 *	-.30
Peer Rejection	-.03 **	.41
Home Aggression of Child	.11 **	-.29
Lack of Social Participation	.06	.28
C. Parent's Aggression (r for mother's aggression = .15; r for father's aggression = .35)		
Home Aggression of Child	-.05 **	.36
Parental Rejection	-.03 *	.32
Spouse's Aggression	.24	.29
Punishment for Dependency	-.10 **	.33
Nonrecognition of Child's Needs	-.37 *	-.02
D. Punishment for Aggression (r mother-father = .42)		
School Aggression	.03	.31
Home Aggression of Child	.47	.46
Confessing by Child	-.17	-.35
Parental Disharmony	.28	.18
Parental Rejection	.20	.44
Parental Restrictiveness	.42	.28
Approval of Home Aggression	.00	-.28
Rural Background	-.26	-.16
Shame at Home	.16	.32
Child's Dependence Avoidance	.00 *	.38
Nonrecognition of Child's Needs	-.35	-.14
E. Confessing by Child (r mother-father = .25)		
School Aggression	-.21	-.31
Peer Rejection	-.10	-.29
Home Aggression of Child	-.28	-.07
Residential Mobility	.04 *	-.32
Parental Rejection	-.07 **	-.58
Approval of Home Aggression	-.32	-.14
Rural Background	.12	.28
Lack of Social Participation	.04 *	-.34
Punishment for Dependency	.10 *	-.28
Punishment for Aggression	-.17	-.35
Punishment for Nurturance	-.04	-.28

* Difference between r for mother and father significant at .10 level.** Difference between r for mother and father significant at .05 level.

there is no significant relation between mother and father). Here it is the father's score that is discriminating. Whereas mother's ethnicity was not significantly related to any other variable, father's ethnicity was significantly related to rejection by peers and home aggression of child, among other things. The relation between father's ethnicity and child's behavior is es-

pecially significant when fathers and sons are paired. For example, father's ethnicity related $+ .53$ ($p < .001$) to son's school aggression and $-.53$ to his home aggression. One possible interpretation would be that ethnic fathers do not tolerate aggression in the home, and thus the tendency on the part of the son to aggress is displaced to the school context. Another possibility is that ethnic fathers may provide a more aggressive role model by their punitive behavior at home. At any rate, it would seem that father's ethnicity has a more important influence on the child's behavior and the reaction of others to him than does mother's ethnicity, at least for boys (correlation between father's ethnicity and daughter's aggression fails to reach significance at the .05 level; indeed, the only significant correlations for ethnicity of girls' fathers are with social participation [$-.40$] and sanctions for nurturance [$-.42$]).

The likelihood of contaminated data which is present when both predictor and criterion measures are obtained from the same informant, as is often done in childrearing studies, is seen in the relation of the parent's self-rating of aggression to other variables. This measure is not related to school aggression or any other independently obtained measures but is related, at least for fathers, to ratings of the child's home aggression, as well as spouse's aggression and rejection. The relation of parent's aggression to rejection is especially marked for father-daughter pairs (.52). Furthermore, those fathers who score high on rejection of their daughters also tend to rate their wives as aggressive (.49). Thus, those respondents who are willing to admit that they themselves are aggressive will also tend to say that wife and daughter are aggressive and that daughter does not measure up to their standards. Since father's self-rating of aggression related only to his own rating of child's aggression and not at all to the independently obtained rating, response set is certainly a compelling consideration.

Again, although both mother's and father's punishments for aggressive behavior are related to the frequency of aggression at home, in each case both these sets of items are reported by the same informant, and response set must also be considered as a reason for this relation. Only father's punishment is significantly related to the independently obtained measure of school aggression. This relation holds especially for father-son pairs ($r = .46$). Thus, it seems that what the father does in response to the child's home aggression is more likely to influence aggressive behavior of the son than what the mother does. The same results obtained when each of the 13 items of aggressive behavior in school which comprised the aggression score from the parent's interview were considered separately. It was still the more punitive father who had the more aggressive child in school. The results for mother were in the same direction, but none of the correlations between punishment by the mother for a specific behavior and the appearance of that behavior in school was significant. The importance of the father as a role model for aggressive behavior is here strongly suggested. Levin and Sears (8) have stressed the importance of the father as a model. In a study of

fantasy aggression they found that boys (attending nursery school) who were punished by their fathers and who were also identified with their fathers tended to be more aggressive in the doll play situation than those not so identified or punished. Their measure of identification was a rating of superego formation based on four open-ended questions having to do with confessing and denying by the child of proscribed acts he was known to have committed. These questions were asked of mothers only.

In the present study two of these four questions were adapted to a closed-end version, thus forming the scale of confessing. They were asked of both mothers and fathers with differing results. It was found that indeed, confessing behavior of the child as rated by the father was related to the independently obtained measure of school aggression, but in the opposite direction from the Levin and Sears results and with these results holding for both boys and girls ($-.39$ for girls; $-.37$ for boys). However, the relation between confessing and school aggression was not significant either for mother-son or mother-daughter pairs. The same holds true for the relation between confessing and punishment for aggression. There is a significant relation for fathers but not for mothers, especially for father-son pairs ($-.53$). The important feature in the aggression-confessing-punishment relation may indeed be the punitiveness which would lend support to the role model hypothesis. At any rate, these results are similar to those of Levin and Sears in stressing the importance of the father, especially for boys. The difference in direction of the results is likely a function of the difference in the nature of the criterion, in one case real-life aggression as rated by the child's peers, and in the second case, fantasy aggression, as observed in doll play by the experimenters. In the latter case, the child, who is rated by the experimenter, is actually playing the role of the "daddy" doll; in the former he is rated by his peers as he, the child himself, behaves in the natural setting. Also accounting for the variation may be the difference in age level, preschool vs. third grade, as well as the different operations used in measuring the variable, confessing.⁵

Discussion thus far has dealt with correlations between measures obtained from the same informant. When cross-over correlations are considered, e.g., the correlation of each mother score with all other father scores, there is considerable shrinkage in size of correlation. For example, the correlation between father's rating of rejection and father's rating of frequency of aggression is $.46$, but the correlation between father's rating of rejection and mother's rating of frequency of aggression is only $.15$.⁶ A separate analysis was also done in this manner for each of the 13 individual items of aggres-

⁵ Levin (7) has stated that the use of superego development as an index to identification, "assumption of adult role models by the child," may be valid for young children but that by the age of 8 years role modeling and superego functioning quite likely have differentiated into separate behavior patterns.

⁶ For an extended discussion of the results of this cross-over analysis, see paper by Banta *et al.* (1).

sive behavior which go to make up the composite child's home aggression score. A significant relation between mother's punishment and mother's rating of child's home aggression was found on five items, between mother's punishment and father's rating of child's home aggression on only two. In the same way, father's rating of punishment correlates with father's rating of child's home aggression significantly on four of the 13 items and with mother's rating of child's home aggression on none of them. This consistent drop, when going from correlations between predictor and criterion obtained from the same individual to correlations between predictor and criterion each independently obtained, reinforces suspicions of contamination in those studies which use only the mother as informant about both child-rearing practices and child behavior.

In general, it would seem that mothers and fathers observe, sanction, and report their children's behavior from different vantage points. Which one should we accept, and which ignore? Traditionally, childrearing studies have utilized information obtained only from the mother. Our results show that indeed mothers may be more discriminating observers in some areas; however, in others, fathers give us more consistent results. As noted above, mother's rejection related to all the classroom measures and none of the interview measures, while the reverse was true for fathers. Father's punishment for aggression and father's ethnicity, on the other hand, related to school aggression, while neither mother's punishment nor ethnicity did. Father's ethnicity was also related to rejection by peers as well as to a number of variables within the interview, while mother's ethnicity was related to no other variable. Moreover, these relations were more marked for father-son than for father-daughter.

It seemed obvious that both fathers' and mothers' reports had to be taken into account. Therefore, an attempt was made to predict from combined mother and father scores on various interview variables to the classroom aggression scores, and indeed it was found that variables which did not relate when predictions were made from the scores of only one parent did show a relation to the criterion when scores of both parents were considered. On each variable the mothers and fathers were split into high and low groups at the median score for their respective samples. The median classroom aggression scores for the four cells containing subjects classified according to whether mother and father were each or both high or low on the specific variable appear in Table 5.

This combination of mothers' and fathers' scores clarifies the relation between father's ethnicity and child's school aggression. If it is only when there is a high ethnic father and low ethnic mother that the aggression score is very much elevated. When both mother and father are ethnics or both are nonethnics, the aggression score is near the median for the entire group. On the other hand, when the mother is high ethnic and the father is low ethnic, the median school aggression score is noticeably low. This suggests that there may indeed be a conflict of cultural values behind the obtained

CHILD DEVELOPMENT

TABLE 5

MEDIAN SCHOOL AGGRESSION SCORES OF SUBJECTS BY COMBINED PARENT POSITION ABOVE AND BELOW MEDIAN ON SOME INTERVIEW VARIABLES

<i>Variable</i>	Hi FATHER		Lo FATHER	
	<i>Hi Mother</i>	<i>Lo Mother</i>	<i>Hi Mother</i>	<i>Lo Mother</i>
Ethnicity	15.5 (14)	69 (8)	6 (9)	16 (19)
Confessing by Child	5 (15)	22.5 (8)	16 (13)	25.5 (14)
Parental Disharmony	15.5 (16)	14 (8)	45 (8)	8.5 (18)
Parental Rejection	36 (17)	3.5 (8)	13 (9)	4.5 (16)
Punishment for Aggression	21.5 (16)	34 (9)	7.5 (10)	4 (15)
Punishment for Dependency	19 (12)	22 (11)	10 (13)	8 (14)
Parent's Aggression	16 (13)	22 (11)	18 (11)	4 (15)
Lack of Social Participation	16 (19)	15.5 (6)	22.5 (6)	5 (19)
Shame at Home	12 (25)	18 (12)	18 (11)	2 (2)
Shame Out of Home	22 (14)	10.5 (12)	25 (10)	10 (14)

NOTE.—Number in parenthesis refers to frequency of subjects in cells (total $N = 50$). Median school aggression score for entire group of 50 children = 15.

ethnicity-aggression relationship. The high ethnic woman, married either to a low or high ethnic man, perhaps brings Old World values into her relationship with him, deferring to him as the decision-maker, authoritarian, etc. Their roles are well defined. On the other hand, the low ethnic woman, married to the high ethnic man, has a different conception of the marital relationship which then conflicts with what is expected by the husband, value conflicts ensue, role definitions are not agreed upon, and the children are deviant, at least as far as peer rating of aggression is concerned.

The importance of getting evaluations from both mother and father is again seen in the relation of confessing to aggression. It is only for those children whose mother and father agree that the child does not deny his guilt that the aggression score tends to be low. When both agree he does deny, the aggression score tends to be high. When there is disagreement between mother and father on this variable, the result is closer to the median for the entire group.

Parental disharmony operates differently. When both mother and father agree that parental disharmony is low, the child's aggression score tends to be low. When both agree parental disharmony is high or when father says it is high and mother says it is low, the child's aggression score tends to be around the median for the whole group. However, when the mother says disharmony is high and father says it is low, the child's aggression score tends to be very high. Thus, the mother's rating seems to be the crucial one here as it does in some aspects of the rejection-aggression relationship. But the father's role in the latter relationship is also important. Both mother and father must be high on rejection for the child's aggression score to be elevated, but only a low score for mother is needed for the child's aggression score to be low. When father is low and mother high, the aggression

score is at the median for the whole group. In other words, there must be two rejecting parents to make for a high aggressive child, but only the mother's lack of rejection results in a low aggressive child. Quite the opposite is true for punishment for aggression. Here it is just when the father is high that the child's aggression score tends to be high. This is especially true when the father is high and the mother is low. When father is low in punishment for aggression, regardless of mother's score, the child tends to be low on aggression. With punishment for dependency, only low punishment by the father is associated with low aggression. When the father is a high punisher for dependency, regardless of whether mother is high or low, the child's aggression score tends to be moderately elevated.

The results with parent's aggression, when data from mother and father are considered jointly, indicate that an erroneous interpretation can be drawn if only one parent is taken into account. The father's importance as a role model for aggression was suggested in the results mentioned above; however, it is only when both mother and father rate themselves as low on aggressive behavior that the child's aggression score tends to be low. When either or both rates himself as high, the child's score tends to be at the median or slightly higher. The same is true for lack of social participation and shame at home; however, for shame out of home it is the low mother only who tends to have the low aggressive child. It is interesting that there are only two families in which both mother and father rate themselves as low on shame at home. In general, however, it should be apparent from this rudimentary analysis that, for a more complete picture of the effect of socialization practices and attitudes on child behavior, the contribution of both mother and father to this effect must be taken into account.

SUMMARY

Sixty subjects were selected out of a pool of 158 children on the basis of scores on an aggression sociometric administered in their classrooms. All of the mothers and 50 of the 58 available fathers of these children were then independently interviewed to gain information on a number of presumed psychological and social antecedents to aggressive behavior. It was found that mothers and fathers did not agree to an appreciable degree in rating either their children's behavior or their interactions with their children. Even on those variables where agreement between mother and father was moderately high, the relation to other variables was not the same for the two groups of parents. There was a consistent shrinkage in size of coefficient when proceeding from correlations between predictor and criterion which were both obtained from the same parent to correlations when the predictor was obtained from one and the criterion from the other. When an outside, independently obtained criterion was used, very often the father's scores related more adequately than did the mother's scores. The most meaningful and theoretically interesting relations were found when predict-

CHILD DEVELOPMENT

ing from a combined mother and father score to the independent criterion. Only recently have studies emphasizing the importance of the father in socialization of the child begun to appear. These results are further evidence of his importance both as a new method and a new dimension in child-rearing research.

REFERENCES

1. BANTA, T. J., WALDER, L. O., & ERON, L. D. Convergent and discriminant validation of a childrearing survey questionnaire. Paper read at East. Psychol. Ass., New York City, April, 1960.
2. BRONFENBRENNER, U. Family structure and development. Presidential address, Div. 7, Amer. Psychol. Ass., Washington, D.C., September, 1958.
3. CHAMPNEY, H. Measurement of parent behavior. *Child Developm.*, 1941, 12, 131-166.
4. CURETON, E. E. Rank-biserial correlation. *Psychometrika*, 1956, 21, 287-290.
5. DAVIDSON, K. S., SARASON, S. B., LIGHTHALL, F. F., WAITE, R. R., & SARNOFF, I. Differences between mothers and fathers of low anxious and high anxious children. *Child Developm.*, 1958, 29, 155-160.
6. GUILFORD, J. P. *Psychometric methods*. (2nd Ed.) McGraw-Hill, 1954.
7. LEVIN, H. Discussion of symposium: some methodological considerations in childrearing research. Paper read at East. Psychol. Ass., New York City, April, 1960.
8. LEVIN, H., & SEARS, R. R. Identification with parents as a determinant of doll play aggression. *Child Developm.*, 1956, 27, 135-153.
9. MILLER, D. R., & SWANSON, G. E. *The changing American parent*. Wiley, 1958.
10. PETERSON, D. R., BECKER, W. C., HELMER, L. A., SHOEMAKER, D. J., & QUAY, H. C. Parental attitudes and child adjustment. *Child Developm.*, 1959, 30, 119-130.
11. SEARS, R. R., MACCOBY, E. E., & LEVIN, H. *Patterns of child rearing*. Row, Peterson, 1957.
12. SEARS, R. R., PINTLER, M. H., & SEARS, P. S. Effects of father separation on preschool children's doll play aggression. *Child Developm.*, 1946, 17, 219-243.
13. SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, P. S. Some childrearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-234.

Q-SORT DESCRIPTIONS OF FIVE-YEAR-OLD CHILDREN BY THEIR PARENTS^{1,2}

GENE R. MEDINNUS*

San Jose State College

Although the plea for theory construction and new conceptualizations is a frequently repeated one in surveys of the parent-child literature, the need for the development of new methodological tools in child research is also recognized. The study reported here is an exploratory attempt to evaluate the use of the general Q methodology with parents of preschool children. The present paper will describe the procedures followed in developing and administering the item pools and the general results obtained with regard to the reliability of the technique and interparent agreement in relation to such first order variables as sex of child and sex of parent. Subsequent reports in this series will include data concerning the relation between interparent agreement and the child's adjustment as well as findings in the area of the consistency of interparent agreement from one measuring instrument to another.

An attempt has been made to overcome some of the shortcomings present in many parent-child researches: nonurban parents from the lower and

* Psychology Department, San Jose State College, San Jose 14, California.

¹ This research was completed under the sponsorship of the Elizabeth McCormick Memorial Fund of Chicago.

² A summary of the present paper was presented at the West. Psychol. Ass., San Jose, California, April, 1960.

CHILD DEVELOPMENT

lower-middle classes participated as subjects, and information was obtained from the fathers as well as from the mothers. Several writers (1, 11) have pointed out the meager amount of parent-child research data which has been obtained from fathers. Sigel (11) stressed the incongruity of this situation in view of the prominent role which the father plays in Freudian psychology. In an analysis of some of the parent-child literature using a role theory framework, Brim (1) noted that mothers outnumbered fathers two to one as sources of information; this he attributed to the greater availability of mothers as research subjects.

In general, the Q-sort technique, described in detail by Stephenson (12), involves a task which requires the subject to sort a large number of descriptive phrases into groups or piles which he feels characterize himself to a greater or lesser degree. Typically the item pool contains over a hundred items which are sorted into 11 to 13 groups ranging from those which the subject feels are "most characteristic of myself" to the group of items at the other end of the continuum labeled "least characteristic of myself." The number of items to be placed in each pile is usually designated in advance. When used in this manner, the method allows for a systematization of an individual's self-perception as well as for comparisons among individuals.

Rogers (9) has used the Q-sort technique to obtain a picture of the client's self-perception before and after therapy, as well as of the client's view of the "ideal self." The method has been used also to compare interviewers' or clinicians' judgments (descriptions) of clients. Such sorters commonly observe that they have insufficient information about the client to make judgments concerning every item. In this respect it would seem that the technique is especially appropriate for use by parents in describing their children since the parent certainly is most intimately acquainted with the characteristics and behavior of his child and is the best qualified informant in this regard.

PROCEDURE

Subjects

The 76 parents, 19 sets of parents of boys and 19 sets of parents of girls, in the present study were all parents of 5-year-olds who were subjects in an investigation of school readiness. The location of the field study was a small midwestern community of approximately 5,000 residents. The 5-year-old sample itself was not completely representative of the entire population of 5-year-olds in the community so far as intelligence and father's occupation are concerned. The mean IQ of the 5-year-olds was 111.6 according to the Stanford-Binet (Form L) intelligence tests administered by the writer. The socioeconomic status of the children's families was approximately evenly distributed between the upper-lower and lower-middle classes according to Warner's Index of Status Characteristics (13). The mean ages of the mothers

and fathers were 32.1 and 34.1 years, respectively, and the corresponding mean years of education were 12.0 and 12.2.

Construction of the Item Pool

In assembling the pools of items to be used, an attempt was made to obtain as complete a coverage as possible of characteristics which pertinently and significantly describe the 5-year-old level. Child development and child psychology texts and literature were consulted, and numerous child behavior rating scales were examined. Overlapping descriptions of behavior were eliminated, and those items were used which showed the greatest frequency of mention in the above sources as well as those thought by clinicians to be most predictive of later adjustment (research evidence on this point is meager, however). In addition, since the larger study dealt with first grade adjustment, several items were taken from a first grade adjustment scale constructed by the writer (8), based on interviews with 25 first grade teachers. In general, characteristics covering the following major areas of child behavior were included: emotional behavior, social behavior, physical status and motor behavior, and intellectual abilities. So far as possible, the items were phrased colloquially in language familiar and comprehensible to parents.

In the preliminary pilot investigation with a small group of parents, positive and negative characteristics were included in a single pool. The parents were asked to sort the items into seven piles, ranging from "most characteristic of my child" to "least characteristic of my child." Bimodal distribution of items resulted in all cases; a majority of the positive items were placed near the "most characteristic" end of the continuum, while most of the negative items fell at the "least characteristic" end. This resulted in little discrimination among the items for the various parent sorters. To obtain such discrimination it was necessary to divide the positive and negative characteristics into separate pools. Additional evidence for the necessity of controlling the social desirability of items is presented in a study by Edwards (3) in which 50 male and 50 female college students made "self-sorts" using 135 items which had been rated previously on a scale of social desirability. Product-moment correlations of .84 and .87 were found between the mean weights assigned the items by the sorters and the measure of social desirability of the items for the male and female students, respectively.

It was decided initially to reduce the number of items customarily used in the Q-sort procedure. One purpose of this was to make the task more acceptable to the parents. Also, research workers using the Q-sort technique often criticize existing item pools on the grounds that they include both phenotypic and genotypic items. Comparisons and judgments must be made between readily observable characteristics and those which involve inferentially-drawn interpretations. Following is an example of each of the two types of items: phenotypic, "Is of average height"; genotypic, "Has an

CHILD DEVELOPMENT

unresolved Oedipus complex." Therefore, the present items were limited to those describing fairly overt types of behavior and/or to those about which parents could make judgments based on day-to-day observation and contact with their children. The number of items in each of the two pools (positive and negative) was limited to 42.

Unforced distributions (the number of items to be placed in each pile not designated) were used in the pilot investigation, but again little item discrimination resulted, since in the sort of the positive pool of items a majority of them were placed at the "most characteristic" end of the distribution while the reverse was true for the negative pool of items. Since it was apparent that a forced distribution was advisable, only the shape of the distribution remained to be determined. The parents in the pilot study were asked to sort the items using both a rectangular and a seminormal distribution. They reported much difficulty in the sort requiring a seminormal distribution of the items; this was due in large part to the very small number of items at either end of the distribution. The parents felt also that the large number of items placed in the middle pile added little to the description of their particular child. This last point is supported by the findings of Livson and Nichols (6), who concluded that the forced rectangular distribution requires the maximum possible number of interitem discriminations. They also present evidence to indicate that "Making more discriminations tends to increase the reliability of the total information carried by the sort" (p. 163).

To obtain a measure of the parents' value system with regard to child behavior and attributes, they were requested also to sort the two pools of items in terms of the "ideal 5-year-old." This is fairly straightforward for the positive pool of items; the items were sorted on a continuum from "most important for the ideal 5-year-old to possess" to "least important for the ideal 5-year-old to possess." Since one would not want an "ideal" 5-year-old to possess any of the characteristics included in the negative pool, the procedure here involved arranging the items on a continuum from "least bad for the ideal 5-year-old to possess" to "worst for the ideal 5-year-old to possess."

In summary, the technique involved two sets of items, one containing only positive traits or characteristics and the other only negative aspects of behavior. Each pool consisted of 42 items typed on separate cards. The cards were sorted into seven piles ranging from "most characteristic of my child" to "least characteristic of my child," with six cards in each pile. Following is a list of the 84 items.

PLUS POOL

1. Can take disappointments calmly
2. Is able to give and accept affection from others

MINUS POOL

1. Can't seem to stick at anything
2. Gets mad easily

GENE R. MEDINNUS

PLUS POOL (*continued*)

3. Is obedient; usually does what he is told to do
4. Girl: very aware of herself as being a girl
Boy: very aware of himself as being a boy
5. Gets over being mad quickly
6. Has a pretty even disposition—doesn't have many ups-and-downs in how he feels
7. Is interested in learning new things; asks many questions
8. Is willing to take care of his own possessions, such as toys
9. Has a lot of energy and pep; doesn't get tired very quickly
10. Gets along fine in new (or strange) situations
11. Catches on to new ideas quickly—for example, is quick at learning how to play new games
12. Likes to be read to
13. Is able to take correction and criticism without getting all upset
14. Can be counted on to carry out a job which he is given to do
15. Is forgiving; doesn't hold grudges
16. Is able to keep himself busy; can figure out for himself what to play or what to do next
17. Has a good imagination
18. Uses good language and words for his age; has a good vocabulary
19. Is in good physical condition; usually healthy
20. Gets along well with other children; is well liked by other children
21. Is sympathetic toward others' hurts and misfortunes
22. Is a self-starter; has a lot of initiative

MINUS POOL (*continued*)

3. Is easily influenced by other children; can be talked into doing things
4. Always thinks he has to be the best when he's doing things with other children
5. Has temper tantrums
6. Is nervous and highstrung
7. Doesn't make friends easily—takes quite awhile to get acquainted with children he doesn't know
8. Is a daydreamer
9. Likes to be the center of the stage; is a show-off
10. Takes his hurts seriously; becomes worried about a slight bruise
11. Usually takes a back seat when playing with other children
12. Lacks drive; no spark
13. Gives up easily without really trying
14. Is very babyish when sick or hurt; likes to get a lot of sympathy from others
15. Seems to be afraid of quite a few things
16. Every little hit from another child calls forth a full battle
17. Is pretty stubborn at times
18. Weak physically; poor health
19. Gets real mad and upset if he can't have his own way
20. Often blames others for things he does
21. Whines or pouts quite a bit
22. Cries easily

CHILD DEVELOPMENT

PLUS POOL (*continued*)

23. Has a good sense of humor; can understand jokes
24. Is cooperative in family affairs
25. Is able to follow directions; has a good memory for what he is told to do
26. Is good looking, an attractive child
27. Is considerate of other people
28. Is fair—plays fair with other children
29. Has self-confidence
30. Is usually happy, cheerful
31. Sticks up for himself; doesn't let others run over him
32. Is willing to confide in others (parents, teacher, etc.)
33. Is usually a leader in the group he plays with
34. Is willing to share his things with others
35. Has a lot of courage; is brave
36. Is wide awake as far as what is going on around him
37. Has a good idea of what he is and isn't capable of doing; doesn't think he can do more than he really can
38. Shows good judgment; is able to profit from experience
39. Is courteous and mannerly
40. Likes to explore and investigate things
41. Knows the difference between right and wrong
42. Seldom picks fights with other children; is not quarrelsome

MINUS POOL (*continued*)

23. Doesn't seem to be able to plan ahead very well
24. Is easily confused in carrying out a simple task; forgets what to do next
25. Has trouble holding onto small objects, such as scissors, jigsaw puzzle pieces, etc.
26. Is not very considerate of the property rights of others
27. Doesn't like to play with other children
28. Is unpredictable; I never know what he's going to do next
29. Is always on the go; restless
30. Has a habit of lying about some things
31. Feelings are easily hurt
32. Doesn't show common sense sometimes
33. Immature—doesn't act his age
34. Is afraid to try new things, such as new games, new play equipment
35. Sulks and mopes around for quite awhile after being punished
36. Doesn't get along too well in a group of children
37. Is impulsive; acts on the spur of the moment
38. Is noisy and loud
39. Is shy and bashful around strangers
40. Is boastful or bragging at times
41. Is sort of a slow mover—takes his time
42. Is a worrier

Method

Individual evening appointments at the convenience of the parents were made in advance; in every case the investigator visited the home for the data collection. The sorting was done independently and simultaneously

by the mother and the father, seated at opposite ends of a table. The directions for sorting were given, and any questions were answered. After each sort was completed, the items in each pile were placed in appropriately marked envelopes. The order of presentation of the item pools was as follows: the positive and negative pools, in that order, were sorted to describe the parents' own 5-year-old (real sort); then the two pools were sorted to describe the ideal 5-year-old (ideal sort). The items in each of the two pools were numbered from 1 to 42. After each set of parents completed the sorting, their placement of the items was recorded, and the items were rearranged in numerical order before the task was presented to the next set of parents. This procedure standardized the order placement of the items so as to eliminate possible effects of differences in interitem context from parent to parent.

Reliability of the sorts, with regard to their stability over time as determined by the test-retest method, was also obtained. Ten sets of parents, five of them parents of girls and five of boys, selected at random, resorted the pools of items four months, plus or minus one week, after their initial sorts.

RESULTS

Feasibility of the Technique

The parent subjects appeared to understand readily the procedure of Q-sorting although none had had previous acquaintance with the technique. Likewise, the phrasing and meaning of the items were comprehended for the most part. The task was approached with interest and performed conscientiously by the parents. Some of the subjects experienced difficulty in deciding upon the placement of certain items, but, in general, the task produced no negative reactions. In fact, many of the parents enjoyed the sorting and several observed that they had never formalized their perceptions of their child before. It can be said, therefore, that the Q-sort technique is a feasible one for the expression of judgments by parents with regard to their children.

Although the exact time required for the sorting of the four item pools was not recorded, it varied from approximately 45 to 90 minutes.

Reliability

The sort-resort reliability figures are tabulated in Appendix B.³ The reliability coefficients were obtained in the following manner: A correlation coefficient was computed between each parent's original sort and his sub-

³ The appendixes (A, B, C) have been deposited as Document number 6719 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C. A copy may be secured by citing the Document number and by remitting \$2.50 for photoprints, or \$1.75 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

CHILD DEVELOPMENT

sequent sort for each of the four sortings (plus and minus pools for the real and ideal sorts). The means of reliability coefficients for the four sortings by the 10 sets of parents were calculated. The reliability coefficients ranged from .01 to .83; the means for the real sort, plus and minus pools, and for the ideal sort, plus and minus pools, were .57, .55, .51, and .53, respectively.

Interparent Coefficients for Real and Ideal Sorts

The correlation coefficients for interparent agreement in their descriptions of their child (real sort) and for the ideal 5-year-old (ideal sort) appear in Appendix C. The mean coefficients for the total group of fathers and mothers, as well as for the parents of girls and the parents of boys, are summarized in the first four columns of Table 1. Concerning these coeff-

TABLE 1
MEAN CORRELATION COEFFICIENTS FOR INTERPARENT AND
INTRAPARENT AGREEMENT

Parents	INTERPARENT AGREEMENT				INTRAPARENT AGREEMENT			
	Real Sort		Ideal Sort		Plus Pool		Minus Pool	
	Plus Pool	Minus Pool	Plus Pool	Minus Pool	Mother	Father	Mother	Father
Of Girls37	.37			.25	.19	.40	.20
Of Boys43	.49			.28	.34	.33	.32
Combined40	.43	.28	.30	.27	.26	.36	.26

cients, it can be noted that (a) the mean interparent correlations are higher for the real sorts than for the ideal sorts, indicating that parents agree more in how they perceive their child than in their expectations and goals for him; and (b) the mean correlations for the parents of boys are higher than for the parents of girls.

Intraparent Coefficients for Real and Ideal Sorts

An individual parent's two sorts (real and ideal) of the two pools of items were correlated. These coefficients would seem to be an index of the degree to which the parent's perception of his child agrees with his description of the ideal 5-year-old and might be interpreted as the extent of parental acceptance of the child. The coefficients are listed in Appendix C and their means are summarized in the last four columns of Table 1. The means are generally low. They are higher for the fathers of boys than for the fathers of girls on both the plus and minus pools. On the minus pool the mean correlations for the mothers of girls are higher than for the mothers of boys, but the trend seems to be in the opposite direction on the plus pool.

Also, for the girls, the mean correlations of the mothers are higher than those of the fathers, but there is no such difference between the parents of boys.

Rank Order and Mean Placement of the Items

The mean placement of each item in the two pools by the 76 parents is given in Appendix A. The mean placement of the items on the real sort for the various subgroups of parents, as well as for the fathers and mothers separately on the ideal sort, is also listed in Appendix A.

TABLE 2
ITEMS DIFFERENTIATING BETWEEN MOTHERS AND FATHERS
OF BOYS ON REAL SORT

Item	MEAN PLACEMENT		p
	Mothers	Fathers	
<i>Plus Pool</i>			
4. Very aware of himself as being a boy	4.3	2.8	< .05
5. Gets over being mad quickly	3.1	4.0	< .05
8. Is willing to take care of his own possessions, such as toys	1.9	2.6	< .05
<i>Minus Pool</i>			
10. Takes his hurts seriously; becomes worried about a slight bruise	2.0	3.6	< .01
13. Gives up easily without really trying	3.5	2.2	< .01
33. Immature—doesn't act his age	2.2	1.4	< .05
39. Is shy and bashful around strangers	2.6	3.8	< .01

NOTE.—To interpret the means in this and the following tables, it is necessary to point out that the piles into which the items were sorted were numbered from "6" (most characteristic of my child) to "0" (least characteristic of my child) on the real sort. For the ideal sorts pile "6" denoted the most desirable end of the continuum while pile "0" was the least desirable end.

The mothers' and fathers' placements of each item in the ideal sorts were summed separately, and rho was computed between the two groups' rank ordering of the items. Rank order coefficients of .84 and .93 were obtained for the plus pool and minus pool, respectively. The significance of the difference in the mean placement of the items by the following parent groups was calculated: the mothers versus the fathers of boys, the mothers versus the fathers of girls, the fathers of boys versus the fathers of girls, and the mothers of girls versus the mothers of boys. The *t* test was used to assess the significance of the difference between the various means. Table 2 gives those items whose mean placement for the boys were significantly different between the fathers and mothers on the real sort. Comparable items for the girls are tabulated in Table 3. Tables 4 and 5 list those items

CHILD DEVELOPMENT

whose mean placement was significantly different for the fathers of boys versus the fathers of girls and for the mothers of boys versus the mothers of girls, respectively.

Those items whose mean placement was significantly different between the fathers and mothers on the ideal sort are given in Table 6.

TABLE 3
ITEMS DIFFERENTIATING BETWEEN MOTHERS AND FATHERS
OF GIRLS ON REAL SORT

Item	MEAN PLACEMENT		p
	Mothers	Fathers	
<i>Plus Pool</i>			
11. Catches on to new ideas quickly—for example, is quick at learning how to play new games	3.4	4.4	< .05
23. Has a good sense of humor; can understand jokes	3.2	2.0	< .05
<i>Minus Pool</i>			
3. Is easily influenced by other children; can be talked into doing things	4.5	3.6	< .05
7. Doesn't make friends easily—takes quite awhile to get acquainted with children he doesn't know	2.8	1.3	< .01
25. Has trouble holding onto small objects, such as scissors, jigsaw puzzle pieces, etc.6	1.8	< .05
38. Is noisy and loud	3.1	4.2	< .05

TABLE 4
ITEMS DIFFERENTIATING BETWEEN FATHERS OF BOYS AND
FATHERS OF GIRLS ON REAL SORT

Item	MEAN PLACEMENT		p
	Fathers of Boys	Fathers of Girls	
Plus Pool			
9. Has a lot of energy and pep; doesn't get tired very quickly	4.9	3.2	< .01
21. Is sympathetic toward others' hurts and misfortunes	2.2	3.5	< .05
Minus Pool			
3. Is easily influenced by other children; can be talked into doing things	5.0	3.6	< .01
12. Lacks drive; no spark8	2.1	< .01
25. Has trouble holding onto small objects, such as scissors, jigsaw puzzle pieces, etc.7	1.8	< .05
27. Doesn't like to play with other children	1.0	2.2	< .05
37. Is impulsive; acts on the spur of the moment	5.0	3.6	< .01

GENE R. MEDINNUS

TABLE 5

ITEMS DIFFERENTIATING BETWEEN MOTHERS OF BOYS AND
MOTHERS OF GIRLS ON REAL SORT

Item	MEAN PLACEMENT		p
	Mothers of Boys	Mothers of Girls	
<i>Plus Pool</i>			
1. Can take disappointments calmly	1.2	3.0	< .01
17. Has a good imagination	4.8	3.4	< .05
40. Likes to explore and investigate things	3.5	1.9	< .05
<i>Minus Pool</i>			
2. Gets mad easily	4.5	3.1	< .01
11. Usually takes a back seat when playing with other children	2.1	3.5	< .05
16. Every little hit from another child calls forth a full battle	3.4	2.1	< .05
38. Is noisy and loud	4.2	3.1	< .05
39. Is shy and bashful around strangers	2.6	4.3	< .05

TABLE 6

ITEMS DIFFERENTIATING BETWEEN MOTHERS AND FATHERS
ON IDEAL SORT

Item	MEAN PLACEMENT		p
	Mothers	Fathers	
<i>Plus Pool</i>			
10. Gets along fine in new (or strange) situations	3.3	2.4	< .05
35. Has a lot of courage; is brave8	1.6	< .05
<i>Minus Pool</i>			
21. Whines or pouts quite a bit	1.6	2.3	< .05
22. Cries easily	2.2	3.1	< .05
37. Is impulsive; acts on the spur of the moment	5.1	4.2	< .05
38. Is noisy and loud	4.2	3.3	< .05

DISCUSSION

Because the Q-sort technique possesses a number of advantages over other methods of obtaining information concerning parents' perception of and attitudes toward their children, its use is indicated in many situations where such information is considered necessary or desirable. The technique

is superior certainly to the method of parental rating of the child on various scales or variables. It has been the writer's experience that with such a method the majority of parents tend to rate their child "average" or above, which, of course, eliminates variation among parents as well as variation from one variable to another for any one parent rater. Consequently, little significant information is gained from this approach. Check lists typically provide even less information. Advantages of the Q-sort technique include: (a) the derivation of an index of agreement between the parents either through correlational methods or through simply a count of the difference in placement of each of the items employed; (b) the derivation of an index of agreement between a single parent's real and ideal sorts; (c) in effect, the placement of each item can be considered as the parent's rating of the child on that variable which also provides comparisons among the various items or characteristics; (d) an unsubstantiated observation by the investigator would indicate that the procedure tends to result in increased communication between parents, especially with regard to the child under consideration. Animated discussions between parents frequently followed the completion of the sorting. The implications of this for clinical approaches to parent counseling are numerous.

While the reliability coefficients are not high, several factors might have tended to reduce them. The time interval of four months, combined with the fluctuating nature of child behavior at this age level, was perhaps an important factor in this regard. A time interval of one to three weeks between sorts probably would have raised the reliability coefficients considerably. That the changeableness of child behavior at the 5-year level was a factor influencing the magnitude of the reliability coefficients was substantiated by the observations of several parents during the resorting when they noted the extent to which their child had changed with regard to specific items of behavior in the months intervening between the two sorts.

The moderate reliability coefficients strongly suggest also that single measures of parent attitudes and of parental perceptions of their children are not adequate. This point has important implications for those investigations which attempt to relate parent attitudes to child behavior and adjustment. Few, if any, studies exist in the research literature in which concern is shown for variations over time in the parent variable under consideration. The writer has stressed elsewhere (7) the necessity for recognizing the possibility of little long-term constancy in various parent attitudinal characteristics; a number of findings were mentioned which indicate the existence of shifts in parent attitudes, parent behavior, and the like. It might be postulated that the nature, extent, and quality of these shifts are as important determiners of child adjustment as are measures of parent attitudes elicited at any one point in time.

Interestingly, the reliability coefficients were no higher for the ideal sorts than for the real sorts. This seems contrary to expectation since one might define the ideal sorts as representing the parents' set of values regard-

ing child behavior as well as the goals which the parents hope to attain through their child training procedures. Both of these would seem to be fairly stable over time. It might be argued, however, that those specific aspects of the child's behavior which are sources of difficulty to the parent at any given time are those which the parent is most likely to consider important. For example, if a child has been ill a great deal during a certain period, the parent might attach great importance to item 19, "Is in good physical condition; usually healthy," when describing the ideal child. This is to say that the day-to-day changes in a child's behavior and the problems he presents influence the parent's description of the ideal child as well as his description of the child himself.

That the interparent agreement coefficients were higher for the real than for the ideal sorts is understandable perhaps in view of the objective referent for the former sort as compared with the latter. It might be postulated also that there is greater interparent communication concerning the actual child, his behavior, and his characteristics than there is in the area of parental expectations, demands, and goals of childrearing.

The trend toward higher interparent correlations for parents of boys than for parents of girls is difficult to interpret. Perhaps boys, by the very nature of their high activity level and personality traits in general, necessitate greater interparent discussion concerning their behavior. With little research on this point to turn to for an explanation, one might assume that the behavior of boys is more obvious, more overt, and less subtle than is true for girls; consequently, this would lead to greater interparent agreement in perception of boys than of girls. In any case, further verification of the present finding would require research designed to sort out the possible causes for this.

Again, with regard to the intraparent correlation coefficients, there are few data in the literature to assist in interpreting the trends found. One investigation (4), using a paper-and-pencil test designed to measure parental acceptance of children, found that mothers scored higher (i.e., greater acceptance) than did fathers. This would lead us to predict higher intraparent coefficients (real-ideal correlations) for mothers than for fathers; such was not found to be the case in the present study, however. Further analysis (2) of the same data revealed no differences between mothers and fathers with regard to acceptance of children of one sex or the other.

Several findings, on the other hand, would lead us to predict higher real-ideal coefficients for mothers of girls than for mothers of boys. In a study of maternal attitudes and practices in child rearing from birth to the age of 2 years, Hubert and Britton (5) found that mothers seemed to enjoy their daughters more than mothers did their sons. They conclude that "Mutual identification between mother and daughter and between father and son is stronger than mother-son and father-daughter identification, at least as seen indirectly by the mothers" (p. 219). In another study in which the data were obtained again through interviews with the mothers, Sears,

Maccoby, and Levin (10) found that in infancy girls were somewhat more warmly treated than were boys. While both of these studies refer specifically to the period of infancy, it is interesting to note that in the present study, data regarding mothers' perceptions of their 5-year-olds do not indicate that mothers accept their daughters more than do mothers accept their sons—that is, if the real-ideal correlation coefficients can be interpreted as indexes of parental acceptance. The coefficients listed in the last four columns of Table 2 would indicate, however, a trend toward higher intraparent coefficients for fathers of boys than for fathers of girls. Also, again on the basis of the present interpretation of the real-ideal correlation coefficients, it would appear that mothers accept their daughters more than do the fathers. These latter two trends are in agreement with the conclusion reached by Hubert and Britton concerning stronger same-sex than opposite-sex parent-child identification.

The rank-order coefficients obtained between the total group of mothers' as compared with the fathers' placement of the items in the ideal sort indicate a high degree of agreement between mothers and fathers concerning the value they place on general areas of behavior as well as on specific items of behavior. Low coefficients resulted, however, from correlating the ideal sorts of the individual pairs of parents. These findings would suggest that, while in general there is little parent sex difference with regard to their evaluation of the importance of various aspects of child behavior, any one set of parents may show a great amount of disagreement. There were some indications also that, as groups, fathers and mothers agreed more on the significance attached to undesirable aspects of behavior than to the more favorable aspects.

With regard to those items whose mean placement on the real sorts was significantly different depending upon the sex of the parent and the sex of the child, it is apparent that in general their number is little greater than that expected by chance. Again, as in the comparison of the mean placement of the items in the ideal sorts by the groups of fathers and mothers, there would seem to be no consistent differences between fathers and mothers in their perception of their children regardless of the sex of the child. So far as the sex of the child is concerned, these data, though dealing with a different aspect of the mother-child relationship, agree with the Sears conclusion, based on interviews with a large sample of mothers concerning various childrearing practices, that "There were surprisingly few dimensions on which the sexes were differently treated" (10, p. 401).

The finding that there is fairly close agreement between groups of fathers and groups of mothers in their perception of their children (real sorts) and in the importance which they attach to various items of child behavior (ideal sorts), but only moderate agreement between individual pairs of parents, would seem to have important implications for role theory in the area of parent-child relations. It is not necessary to be concerned with differences between mothers in general and fathers in general in their

role prescriptions for the child since such differences are small. More importantly, it is necessary to describe the differences existing for a specific set of parents; it is this difference which would affect most crucially the child's behavior and adjustment (assuming that there is a relation between these variables).

Although numerous discussions of role theory have included comments concerning the role conflicts produced by varying role demands and role prescriptions which an individual must satisfy, there are few research data to support this notion. Specifically, within the role theory framework, it would be predicted that great discrepancy between the father and the mother in their role prescriptions for child behavior would lead to conflict and subsequent maladjustment on the part of the child. It might be postulated, on the other hand, that a certain amount of discrepancy in role demands is useful insofar as it provides the child with practice and experience in role playing. The corollary of this would be the hypothesis that there is a curvilinear relation between the amount of discrepancy between the parents in their role prescriptions for the child and the child's consequent conflict and maladjustment. With too little practice in role playing in the home, the child might well experience difficulty outside of the home when required to meet the demands of other significant adults. When there is too large a discrepancy between the parents in their role prescriptions, conflict is produced in the child in two ways: first, the child is unable to meet the role prescriptions of both parents, especially when the prescriptions demand opposing types of behaviors; second, the child faces the problem of determining exactly when each of the two roles is required of him.

In a role theory context, Brim (1) pointed out the lack of specificity of the research data with regard to the sex of the parent and the sex of the child. The present findings would suggest that this is not so important as concern with role prescriptions and performances specific to any given family. It is the writer's contention, then, that in general few differences exist in attitudes and behavior as a function of the sex of the parent and the sex of the child. Research in the area of parent-child relations needs to be concerned, rather, with the differences which obtain for one set of parents and in relation to their offspring and in relation to their specific family constellation. Idiographic rather than nomothetic data seem to be indicated.

SUMMARY

The purpose of the present study was to evaluate the use of the general Q methodology with parents of preschool children and to explore the extent of interparent Q-sort agreement in the area of parental perception of child behavior.

The sample consisted of 76 parents, 19 sets of parents of 5-year-old boys and 19 sets of parents of 5-year-old girls. The mean IQ of the children was 111.6. The socioeconomic status of the families was divided approximately

CHILD DEVELOPMENT

equally between the upper-lower and lower-middle classes according to Warner's Index of Status Characteristics. The mean ages of the fathers and mothers were 32.1 and 34.1, and mean years of education 12.0 and 12.2, respectively.

Both parents of each set sorted, independently, two pools of 42 items each descriptive of the behavior of 5-year-olds. One pool contained only positive items (plus pool) and one contained negative items only (minus pool). The items were sorted into seven piles, with six items per pile, ranging from "most characteristic of my child" to "least characteristic of my child." In addition to this "real" sort, the items were also sorted by the parents to describe the "ideal" 5-year-old. Product-moment correlation coefficients were computed between each set of parents on the two sorts (interparent agreement) and between each parent's two sorts (intraparent agreement). Reliability figures were obtained from 10 sets of parents who resorted the items with a four-month interval between sorts.

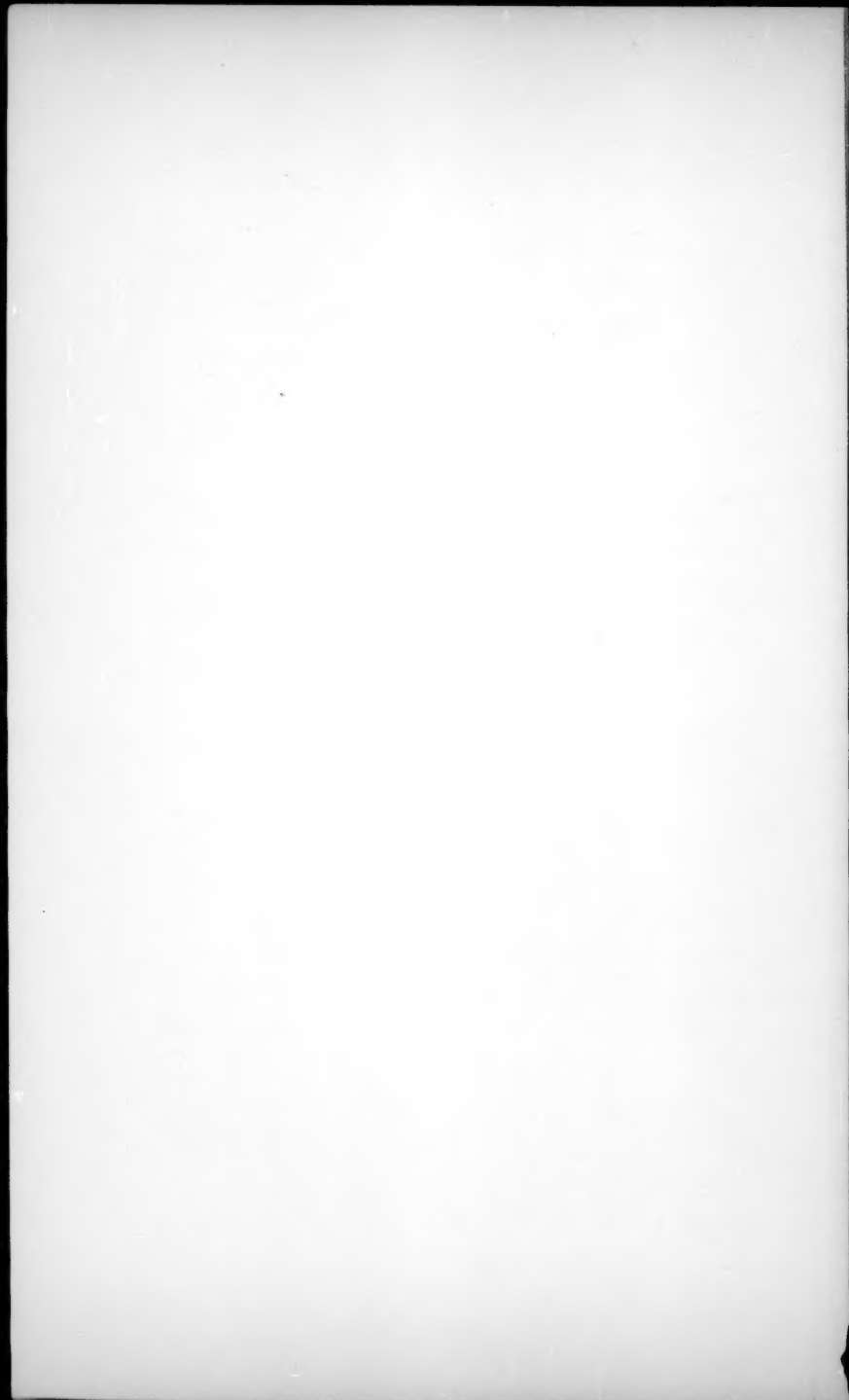
The results and conclusions may be summarized as follows: (a) The mean interparent correlations were higher for the real sorts than for the ideal sorts. This would indicate that parents agree more in how they perceive their child than in their behavioral goals (role prescriptions) for him. (b) If the intraparent coefficients (real-ideal sorts) can be interpreted as the degree to which the parent accepts the child's behavior, it appears that there is only moderate acceptance of the child's behavior among the present sample of parents and that there is wide variability among parents in this regard. (c) Although the individual sets of parents did not agree closely in their ideal sorts, the mean placement of the items by the total group of mothers versus fathers showed high agreement. (d) Few significant differences in the mean placement of items were found between various subgroups of parents (e.g., fathers of girls versus fathers of boys) in their real sorts. That there were few differences as a function of the sex of the parent and the sex of the child was discussed in relation to role theory. The necessity for an idiographic approach in the parent-child area was suggested.

REFERENCES

1. BRIM, O. G., JR. The parent-child relation as a social system: I. Parent and child roles. *Child Develpm.*, 1957, 28, 343-364.
2. BURCHINAL, L. G. Mothers' and fathers' differences in parental acceptance of children for controlled comparisons based on parental and family characteristics. *J. genet. Psychol.*, 1958, 92, 103-110.
3. EDWARDS, A. I. Social desirability and Q-sorts. *J. consult. Psychol.*, 1955, 19, 462.
4. HAWKES, G. R., BURCHINAL, L. G., GARDNER, B., & PORTER, B. M. Parents' acceptance of their children. *J. Home Econ.*, 1956, 48, 195-200.

GENE R. MEDINNUS

5. HUBERT, M. A. G., & BRITTON, J. H. Attitudes and practices of mothers rearing their children from birth to the age of two years. *J. Home Econ.*, 1957, 49, 208-223.
6. LIVSON, N. H., & NICHOLS, T. F. Discrimination and reliability in Q sort personality descriptions. *J. abnorm. soc. Psychol.*, 1956, 52, 159-165.
7. MEDINNUS, G. R. Research implications of several parent-child concepts. *Marriage Fam. Living*, 1959, 21, 329-333.
8. MEDINNUS, G. R. The development of a first grade adjustment scale. *J. exp. Educ.*, in press.
9. ROGERS, C. *Client-centered therapy*. Houghton Mifflin, 1951.
10. SEARS, R. R., MACCOBY, E. E., & LEVIN, H. *Patterns of child rearing*. Row, Peterson, 1957.
11. SIGEL, I. E. The need for conceptualization in research on child development. *Child Developm.*, 1956, 27, 241-252.
12. STEPHENSON, W. *The study of behavior*. Univer. of Chicago Press, 1935.
13. WARNER, W. L., MEEKER, M., & EELS, K. *Social class in America*. Science Research Associates, 1949.



THE RELIABILITY OF MOTHERS' HISTORIES

CHARLES WENAR*

University of Pennsylvania School of Medicine

In view of the extensive use made of mothers' developmental histories of their children, surprisingly little is known concerning their reliability. Clinicians and experimentalists are aware of the fact that reconstructive data distort reality, but there are few studies of the nature and extent of this distortion. Certainly Macfarlane's findings are sobering: In the Berkeley longitudinal study she found that, by the time the infant was only 21 months old, "... the retrospective account (of the physical condition during pregnancy) was so unreliable that we have had to disregard it. . . . Weight at birth was reliably reported. The use of instruments was unreliably reported—only two-thirds of the mothers delivered with instruments reported this fact. The duration of labor showed an average discrepancy of three to five hours. . . . Illnesses, unless outstanding, were frequently forgotten" (6, p. 127). She concluded, "Where retrospective interview data are the only type available, the above findings should limit the over-optimistic use of them as factual" (6, p. 29).

* Department of Psychiatry, Hospital of the University of Pennsylvania, 34th and Spruce Streets, Philadelphia 4.

CHILD DEVELOPMENT

Since the interview shows every sign of remaining a popular research instrument and since many data concerning child development and child-rearing practices are derived from mothers' reports, a methodological enquiry is appropriate. The purpose of this paper is to present a conceptual framework in which to study the validity and reliability of mothers' histories. As will be seen, the spadework for many of the methodological studies has been done in other contexts, and the studies themselves will reveal a good deal not only about reconstructive material but about the parent-child relationship as well.

The schema can best be introduced by fantasizing what the Ideal Research Mother would be like. Such a mother could (a) faithfully record events, feelings, and interpersonal relations as they happen; (b) remember with perfect fidelity; and (c) have the power of total recall upon request. What the offspring of the IRM would be like need not concern us here, since she was created solely to illustrate three essential aspects of studying historical material. The first raises the question of the validity of the mother as an observer, while the second and third are concerned with her reliability or consistency over time.

VALIDITY OF INITIAL OBSERVATION

As a rule, a mother's history can be no more accurate than her original observations and evaluations of events, feelings, and interactions. (An exception to this rule will be discussed in the next section.) Yet, very little is known about this crucial and fascinating area. Arguing from certain clinical and experimental data, one could say the mother is the ideal observer since she is the person most involved and in most frequent and intimate contact with her child. Other kinds of clinical observations and research data could be cited to show that, by her very involvement, a mother is the person most vulnerable to distortions, denials, and the sundry tricks affect plays on cognition.

In its simplest form, determining the mother's validity is a matter of correlating her report with an independent measure of the event. For some kinds of factual information, criteria already exist, and there is reason to be optimistic about more becoming available in the near future. Recent advances in measuring infant responses and the increasing number of normative studies will aid appreciably in solving the criterion problem. For instance, if a mother says her infant is "very restless," this statement can be compared with measures of the infant's gross motor activity and normative data on a large group of infants. Other norms can be used to determine how accurate a mother is in her statements concerning the developmental status of her child or the timing and nature of her childrearing practices.

Attitudinal and evaluative statements are more difficult to validate. Unfortunately, such statements are heavily weighted in many clinical situ-

ations and research projects. When a mother says, "I don't pay any attention to his lisping," or "He never seems to want to be loved," her report may be extremely important in studying the etiology of speech defects or withdrawal tendencies, while being extremely difficult to validate.

Even in this area, however, there are encouraging signs. Smith (9) observed manifestations of dependency in the mother-child interaction and later interviewed the mothers concerning their techniques for handling dependency. She found that 70 per cent of the mothers reported using techniques they were observed to use and that ratings on techniques from the interview were positively related to the total observed techniques at the .05 level. She further found that mothers reported more accurately their use of negative direction than they did their use of positive direction. This was interpreted as meaning they were more conscious of actions which they considered bad. In other observational studies, it would be comparatively simple to ask the mother, at the end of the experiment, to rate the same or similar variables used by the judges. In this way, data can be accumulated on the kinds of behavior mothers can evaluate with a high degree of accuracy and the amount of error which is to be expected.

In such studies, mothers' evaluations are compared with those of judges. Although the latter are not as scientifically pleasing as the reading of a meter, the assumption is that the expert can rate the behavior and attitudes more objectively than the mother. A thoroughgoing skeptic might point out that experts can be victimized by their feelings or their social class or their hypotheses just as the mother can be blinded by her neurosis. Until the day a purely objective science of child development is devised, however, the field can do no better than rely on its experts, standardizing them and controlling them to the best of its ability.

As was mentioned above, such methodological studies can also advance our understanding of the mother-child relationship. Two examples will be briefly described, one concerning "invalid," the other concerning "valid" mothers.

Mothers who are poor observers, instead of being dismissed as sources of error, comprise a particularly interesting subgroup. In clinical settings, a mother's distortions are especially valued as furnishing keys to understanding her personality, her past experience, her unconscious attitudes toward the child and, consequently, her faulty childrearing practices. When a mother calls her child a "terrible liar," but can cite only a few instances in which incomplete homework was claimed to be done, this distortion illuminates both the nature of her value system and the emotional impact of her morality training on her child. Further research could proceed in a number of directions. One is to collect data on the mother's personality development in order to test the clinical hypothesis that present distortions arise from her own childhood conflicts in the same area. Thus, one could hypothesize that the mother who unrealistically complains that her infant is helpless and demanding has herself been inadequately loved

CHILD DEVELOPMENT

as a child and thus is unduly distressed by meeting dependency needs; or the mother who sees her child as always lying has been made to feel guilty about minor transgressions as a youngster. Another line of research is to follow the consequences of the mother's distortions in the child's development. At a simple level there may be a direct relation between distortion and development; the healthy infant who is seen as "always sick" by his mother eventually might become sickly, or an infant who is seen as "mean from birth" may, in fact, become mean. At a more complex level, the child might develop the exact opposite characteristics the parent professes since he responds to strong covert or unconscious needs. In light of the richness of clinical evidence, there is a surprising lack of experimental investigation of distorted perceptions, their relations to other aspects of the mother's personality and history, and their effects on the child's development. In such research, the nondistorting mothers can serve as a ready-made control group.

In regard to "valid" mothers, there is evidence (3) which suggests that they are not necessarily "good" mothers. A mother can correlate highly with experts in evaluating early development, she can also evidence a good intellectual awareness of her child's current thinking, and yet behave in a damaging manner, partly because of a lack of positive feeling and partly because she uses her understanding to manipulate the child so as to achieve her own ends. One can speculate that, eventually, there will be evidence of "invalid" mothers who are genuinely maternal. These would be women who function at an intuitive, affective level and who have little interest in the kind of conceptualization and reflection necessary to evaluate themselves and their children. Certainly, we need to know more about the relation between "awareness" and maternal behavior since the association does not appear to be a simple one.

FIDELITY OF MEMORY

Just as in the previous section it was important to understand the objective situation and the mother-as-perceiver, it is important here to understand the subsequent history of the child and the nature of the mother-as-rememberer.

Subsequent events might be conceptualized as interpolated material and, as is well known, it is the nature of this material rather than the passage of time per se which crucially affects recall. However, it is difficult to hypothesize what the precise nature of its influence will be. Take an hypothetical example of two 4-year-olds with a similar history who attend preschool for a year. One adjusts well, makes friends, discovers new skills, and generally sets up positive reverberations within the family; the other one has separation anxieties, becomes a behavior problem, and increases tension within the family. What will be the differential effect upon recall of early development? Will one mother be set to remember good things, and the

other bad? Will continuity between the immediate and remote past decrease accuracy of recall because of the relative homogeneity of events (i.e., everything going well or everything going badly), while discontinuity heightens accuracy through contrast?

Such questions are difficult enough assuming a kind of normal memory in all mothers. However, the mother's particular intellectual and personality characteristics introduce another set of variables. One would expect a compulsive individual who keeps a baby book and compares her child's development with Gesell's norms to withstand the inroads of all kinds of subsequent events better than an hysterical one who lives from day to day and relies on repression to protect her from painful memories. Paul's experiments (7), which are conceptualized in terms of "cognitive style," also furnish clues regarding different kinds of recall. Using a Bartlett-type situation, he distinguished individuals who import new material from ones who skeletonize material when they recall and further differentiated importations which make the material more coherent from those which merely embellish it. Although such findings are somewhat limited at present, the approach promises to be a valuable supplement to our knowledge of the effect of defense mechanisms and personality disorders upon memory.

The case in which the mother's subsequent experiences increase the validity of her evaluation of early events is of special interest. Change does not have to be in the direction of repression, distortion, and defensiveness, but it can be in the direction of greater clarity of understanding. Having more children, for example, can make a mother realize that her first born was not as fretful and demanding as she had thought. One might speculate that this capacity to reevaluate the past more realistically is one measure of the mother's capacity for psychological growth.

Only a series of empirical studies can determine which of the above variables is significant and in what way they affect memory. Some of these studies can be incorporated into the design of longitudinal projects with relative ease. Macfarlane, in the previously cited investigation (6), has made a start in this direction. Her finding that mothers tend to err in the direction of precocity suggests that, in general, distance lends enchantment to the neonatal period. However, her reference to the complex nature of the data hints that there may be a number of specific relations which must be teased out.

It also may be that the time span studied can be a relatively short one rather than the entire period of childhood. If forgetting here follows the same course that it does in other memory research, about as much information might be lost after two years as after ten. The author is presently conducting a study which compares the histories of the first three years of life given by mothers of 5-year-old and 9-year-old children. Factual information comes from detailed medical records which have been kept since birth on the children. Taking a cue from Macfarlane and the experimental literature, it is hypothesized that there will be only a slight increase in unreliability

CHILD DEVELOPMENT

in the 9-year-old group as compared with the 5-year-old. The data will be further analyzed to ascertain what kinds of information are reliably and unreliably remembered and what kinds of mothers are particularly accurate and inaccurate in their recall.

RECALL UPON REQUEST

The crowning achievement of the Ideal Research Mother is that the total past is totally available upon request. With more fallible women, recall and reporting of information is, to a certain extent, a function of the method used in trying to obtain it.

In comparison with the extensive literature on the art of interviewing parents, the number of systematic studies is relatively small. Fortunately, the technique is shared by a number of disciplines, and there is a substantial body of research emanating especially from social psychology. Many controversies are not settled, but important variables have been isolated and studied, and the psychologist has at his disposal valuable orienting ideas as to the consequences of choosing one technique over the other. For instance, the advantages and disadvantages of the structured and unstructured interview have been spelled out in sufficient detail to allow the interviewer to judge which is more appropriate for his purposes and conduct systematic studies of the two procedures in eliciting historical material. Antonovsky (2) used structured interviews, unstructured interviews, and observation to test a number of hypotheses about childrearing practices and the behavior of children. Of the three techniques, the structured interview was the most satisfactory in that it yielded more correlations in line with theoretical predictions. Further analysis revealed that, contrary to prediction, there was no closer relation between the two sets of interview data than between interviews and observations. In order to account for this, Antonovsky hypothesized that the wide variability on the three techniques which was noted in certain mothers was due to greater anxiety and conflict in them. Again, this hypothesis was not supported. Although such a study raises more questions than it answers, it represents a rare attempt to evaluate a technique and to discover the interaction between techniques and personality variables.

There is also increasing interest in and information about the interpersonal aspects of obtaining information. The mother can no longer be regarded as an eager, passive recipient, waiting to answer all questions to the best of her ability. Instead, she has her own concern about the picture of herself and her child which she presents to the interviewer, as well as her own reactions to his personality. There is ample evidence that the pull toward social acceptability is great, and this problem plagues interviewers and inventory makers alike. The mother's view of the purpose of the interview may be another selective factor, and a developmental history of a child with an ulcer may differ depending on whether it is given to a physician or a social worker.

CHARLES WENAR

Finally, the individual characteristics the interviewer brings to his role probably influences the response of the mother. Again drawing from research in other fields, there is evidence that his class, race, sex, and status significantly affect the kinds of attitudes expressed. Furthermore, respondents tend to say they share their interrogator's views on politics, race relations, war, and prefabricated houses. A similar phenomenon may exist in child psychology, where the interviewer's anxieties about specific areas might subtly distort his questioning and recording or cue a defensive reaction on the part of the mother. Fortunately, there is also evidence that such biases in the data can be controlled by training.

To answer the questions raised in this brief outline would take a major research effort. One might even wryly conclude that, by the time we have sufficient information to validate the mothers' histories, we will know enough about child development to eliminate her as a source of information. This may be so, and the reconstructive approach may eventually wither away or become a technique of convenience rather than a source of basic data. In the meantime, there is no reason to proceed quite as blindly as we have been. The individual studies suggested above will give a clearer idea of what kinds of events can be validly observed and remembered or what kinds of mothers give valid and reliable histories.

THE FACTUAL VS. THE SUBJECTIVE APPROACH

The discussion so far has centered on the question, Is what the mother reports really true? It can be argued that the question itself is irrelevant. Certain personality theories and schools of psychotherapy maintain that it is not reality which matters but the individual's perception of reality. Thus, when a mother says, "The baby was always crying," what is important is the mother's evaluation of the infant as a constant crier and the attitudes of exasperation or desperation which accompany such a perception. It could further be argued that it is just such perceptions, with their distinctive cognitive and affective components, which are the crucial formative influences in the child's life and most worthy of study.

This is no place to discuss the merits of the subjective approach, but the methodological problems it involves should be touched upon briefly.

In one sense, such problems are no different from those of the factual approach. In regard to reliability, the question, Did this really happen? becomes, Did the mother really perceive the event the way she said she did? It is reasonable to assume that evaluations of situations change with time and that an experience takes on a different affective coloring. A pilot study being conducted by the author indicates that recall of feeling is much less reliable than recall of fact; e.g., a mother usually does not change her story about whether she breast or bottle fed her infant, but the passage of time significantly affects her report of how she felt about this activity. If sub-

CHILD DEVELOPMENT

stantiated, this finding would raise serious doubts about relying exclusively on subjective material to reconstruct the significant influences in the child's development.

In regard to validity, it seems likely that the subjectivist-researcher will not be able to dismiss reality so readily as his clinical counterpart. One of the most baffling questions in studying the etiology of childhood psychosis, for example, is whether the mother's deviant attitudes cause the pathology or represent reactions to a constitutionally deviant development. One wonders what kinds of lawfulness would be discovered if mothers were grouped together only on the basis of the similarity of their subjective reactions; it would be surprising if the realistically and the neurotically determined perceptions had the same effects on the child's development.

There are two ways for the subjectivist to avoid the methodological problems under discussion. One is to take an ahistorical tack and claim that it is only the mother's present report, only her present perception of her perceptions, which is important. The past is relevant only as it influences the present. This is a defensible position and one suitable to the study of current parent-child relations. However, it automatically eliminates the mother's history as data on which to reconstruct her child's early life.

The other solution is to leave the reconstruction in the hands of experts. These specially qualified individuals could decide whether the mother is an accurate and truthful person, her areas of distortion, how well her report reveals her true feelings, and how much it is a defense against underlying attitudes. This may settle one set of problems—namely, those concerned with, What are the facts?—but it raises a host of others concerning the nature and reliability of expert judgment. At this point, it becomes a matter of opinion as to whether the cure is worse than the disease.

FRUITFULNESS AND CRITERION VALIDITY

Perhaps, one reason why there have been so few criterion validity studies of the interview is that it has been so fruitful. The supreme example is Freud's use of such material to evolve theories and hypotheses which have supplied a generation of developmental psychologists with research ideas. Or, if one would object to regarding a psychoanalysis as merely a series of interviews, the more modest work of Levy (5) on maternal overprotection or Sears (8) on patterns of childrearing can be cited. As long as interview material reveals consistent relationships within itself (e.g., severe toilet training by cold mothers increases the amount of upset in the child), as long as reported antecedents are consistently found for certain behaviors (e.g., overprotected children have a history of excessive contact with the mother during infancy), and as long as such consistencies either test hypotheses or lead to further hypotheses, the interview justifies its extensive use.

There certainly can be no quarrel with such a defense. Yet it is also true that, without at least some of the methodological studies suggested above,

one can never be sure of the meaning of either positive or negative findings. For instance, Gerard and Siegel (4), using interviews of schizophrenic and normal controls, found a number of significant differences but, what is more to the point, also a number of nonsignificant differences in areas where they would be expected, such as punitiveness of the parent, feeding disturbances, and frequency of behavior problems. At this point, the investigator is hopeless to know whether negative results were obtained because in fact no differences exist between the two groups or because the interview was an insensitive or an inappropriate technique for revealing such differences. Another case in point: Altman (1) found no correlation between judges' ratings of maternal attitudes based on interview material and certain personality characteristics of normal children. However, when the interviewer rated the mothers on the same characteristics, positive relations were found. Thus, there is a suggestion that the interviewer makes the most sensitive judge, since the traditional method of evaluating the data obscured rather than revealed important differences. The more general point is that negative findings will continue to have an ephemeral and arbitrary quality until we are more certain of the nature of our measuring instrument.

The same point holds for positive findings. At present, there is no way of being sure whether the obtained relations are between actual historical events, between perceptions of events, or between perceptions of events and reality. For instance, a mother's memory of her feelings about breast feeding may be completely distorted and yet correlate highly with behavior deviations in her child if a common personality dynamic is responsible for both. At the methodological level, the interviewer who knows the kinds of questions which are likely to elicit unreliable answers and the types of mothers who are likely to distort developmental data will not have his faith in his technique shaken by evidence that reported events did not really happen. The more we can distinguish those parts of a history which represent accurate reporting from those which represent distortions, the more certain we can be about the nature of the relations which are found and the more satisfying the interview will become as a research instrument.

REFERENCES

1. ALTMAN, C. H. Relationships between maternal attitudes and child personality structure. *Amer. J. Orthopsychiat.*, 1958, 28, 160-169.
2. ANTONOVSKY, H. F. A contribution to research in the area of the mother-child relationship. *Child Developm.*, 1959, 30, 37-51.
3. GARNER, A. M., & WENAR, C. *The mother-child interaction in psychosomatic disorders*. Univer. of Illinois Press, 1959.
4. GERARD, D. L., & SIEGEL, J. The family background of schizophrenia. *Psychiat. Quart.*, 1950, 24, 47-73.
5. LEVY, D. M. *Maternal overprotection*. Columbia Univer. Press, 1943.

CHILD DEVELOPMENT

6. MACFARLANE, J. W. Studies in child guidance. I. Methodology of data collection and organization. *Monogr. Soc. Res. Child Developm.*, 1938, 3, No. 6.
7. PAUL, I. H. Studies in remembering. *Psychol. Issues*, 1959, 1, No. 2.
8. SEARS, R. R., MACCOBY, E. E., & LEVIN, H. *Patterns of child rearing*. Row, Peterson, 1957.
9. SMITH, H. T. A comparison of interview and observation measures of mother behavior. *J. abnorm. soc. Psychol.*, 1958, 57, 278-282.

FAMILY RELATIONS OF BRIGHT HIGH-ACHIEVING AND UNDER-ACHIEVING HIGH SCHOOL BOYS¹

WILLIAM R. MORROW*

Fulton (Missouri) State Hospital

and ROBERT C. WILSON

Portland (Oregon) State College

The present article reports data on the family relations of bright high school boys making good grades as compared with bright high school boys making mediocre or poor grades.

Of numerous studies of school achievement, few have obtained data regarding family relations correlates. These few have indicated a positive association between student achievement and emotionally supportive home situations. This general finding applies to elementary school pupils (4, 9, 15, 18, 20,), high school students (13, 14), and college students (12).

More specifically, parents of high-achievers have been found to give their children more praise and approval (18), to show more interest and understanding (19), to be closer to their children (14), to make their children feel more family "belongingness" (20) and identification with parents (19). On the other hand, parents of under-achievers have been reported to be more domineering (12, 14) and overrestrictive (18) and to use more severe and frequent punishment (4, 14), which is at the same time less effectual (4). Parents of under-achievers have also been found more likely either to baby their youngsters or to push them excessively (9) and to present to their

* Missouri State Hospital No. 1, Fulton.

¹ The research reported here was part of a larger study of bright high- and under-achieving high school boys, sponsored by the Gifted Child Project of the Portland, Oregon, Public Schools and supported by a grant from the Fund for the Advancement of Education.

CHILD DEVELOPMENT

youngsters either low or extremely high (pressuring) demands for achievement (18). Finally, homes of under-achievers are reported to show more tension (12) and more parental disagreement as to standards of behavior expected of their youngsters (19).

Most of these studies have been based on qualitative data such as interviews. Many had sampling limitations such as a very small sample and questionable equation of groups. No previous study seems to have dealt with family relations correlates of school achievement among high school students of superior intelligence.

In the present research the groups compared were carefully equated, the sample was relatively homogeneous in intelligence, all subjects being of superior intelligence, and the data were obtained by structured measuring instruments.

Before the present study was initiated, exploratory interviews covering various topics including family relations were held with a few bright high- and under-achieving high school boys.² On the basis of these interviews as well as the studies cited, it was hypothesized that the reported family relations of high-achievers, as contrasted with those of under-achievers, are characterized by: (a) more emotionally supportive home environments; and, more specifically, (b) greater family sharing in recreation, decision-making, and exchange of confidences and ideas; (c) greater mutual parent-child affection, acceptance, trust, and approval; (d) less parental domination, severity, and restrictiveness; (e) more sympathetic parental encouragement of achievement, but less overinsistence on achievement; (f) greater harmony between parents and more regularity of home routines.

It was hypothesized further that students' family relations influence their school achievement via certain mediating variables. Specifically, it was hypothesized that supportive home environments are associated with positive student attitudes toward (a) teachers, (b) school, and (c) intellectual activities, as representing the adult world of parents, and that each of these mediating attitude variables would be associated with school achievement.

METHOD

Sample

The sample included two equated groups each containing 48 high school boys of superior intelligence (120 IQ or above). The groups were equated for grade in school, socioeconomic status,³ and intelligence. Each group

² These interviews were conducted by Morrow, using a schedule which he prepared.

³ After all other criteria had been met except equating the groups for socioeconomic status and IQ, the high-achieving group included five more upper class boys than did the under-achieving group, 12 more upper-middle class boys, and nine more lower-middle class boys. The under-achievers included 12 more lower class boys than did the high-achieving group. This imbalance is consistent with previous findings of a positive association between academic achievement and family socioeconomic status for elementary and high school students (2, 3, 5, 16, 17), though not for college students, for whom findings on this point have been inconsistent [see, for example, (8)].

contained 19 ninth-graders, 14 tenth-graders, and 15 eleventh-graders. Each group likewise contained one upper class, 15 upper-middle class, 23 lower-middle class, and 9 lower class boys. The mean IQ of the high-achieving group was 126.0; that of the under-achievers, 125.3, the difference being nonsignificant.

The groups differed in grade-point average in academic courses (with a minimum of three academic courses for each student). The high-achievers maintained an average of 1.00 to 1.67 (1 being the best grade possible, 5 the poorest) during the school year 1955-1956. The under-achievers maintained an average of 2.75 to 5.00, 2.75 being regarded as under-achievement for students of superior intelligence.

Data Collection

Several types of information about family relations were obtained from the students through group-administered questionnaires. The students were not asked to sign their names, but birthdates and other data were used to identify questionnaires for the purpose of selecting the sample. Otherwise, anonymity was maintained and information on individuals kept confidential.

The students' family relations as seen by themselves were evaluated primarily by 16 self-report Family Relations Scales. Each scale consisted of six questions about the student's relations with his parents (or foster parents). The scales were presented in consecutive order, but without scale titles or breaks in spacing. Each student was asked to indicate to what extent each item described his own home situation, using the following four response categories:

- | | | |
|-----------------|----|-------------------------------|
| 1. Not at all | or | Almost Never |
| 2. A Little | or | Sometimes |
| 3. Considerably | or | Often |
| 4. Very Much | or | Very Often (or Almost Always) |

The student was assigned a score on each scale by summing his scores on the six items in the scale. A score was also obtained on the total of all scales (except Harmony of Parents⁴), conceived as an index of Over-all Family Morale.

The general form of the scales, as well as much of their content, was adapted from a questionnaire developed by Brown and his associates (10, Ch. 21).⁵ Seven scales were added; and a number of Brown's items were omitted, revised, or supplemented by additional items. Other sources of item

⁴ The Harmony of Parents Scale was scored only for students with two parents living together. The *N* for this scale was 40 in each achievement group, re-equated for grade in school, socioeconomic status, and intelligence.

⁵ We wish to thank Dr. Robert J. Havighurst for making available to us a copy of the Family Relations Questionnaire used in his research. That questionnaire was developed by Dr. Andrew Brown and his associates for use in a larger study of adolescent character and development, directed by Dr. Havighurst and Dr. Hilda Taba.

CHILD DEVELOPMENT

content included instruments developed by Block (1), Itkin (11), Rickard (18), Stott,⁶ and Strodtbeck,⁷ as well as our preliminary interviews.

Following are the scale titles and a sample item for each scale:

1. *Family Sharing of Recreation.* "Do you and your family go on picnics or outings or trips together?"
2. *Family Sharing of Confidences and Ideas.* "Do your parents discuss their work and activities with you?"
3. *Family Sharing in Making Decisions.* Do your parents let you help decide everyday family policies, rules, and ways of living?"
4. *Parental Approval.* "Does either parent ever seem to wish that you were a different sort of person?" (Negative item, scored in reverse.)
5. *Parental Affection.* "Do your parents openly show affection for you by word or action?"
6. *Parental Trust.* "How confident do your parents seem to be that you will behave properly away from home?"
7. *Parental Approval of Peer Activities.* "Do your parents object to some of your activities with your friends and acquaintances?" (Negative item.)
8. *Student Acceptance of Parental Standards.* "Do you agree with your parents' ideas about life?"
9. *Student Affection and Respect for Parents.* "Would you like to be the same kind of parent that your parents have been?"
10. *Lack of Parental Overrestrictiveness.* "Do your parents try to direct your activities?" (Negative item.)
11. *Lack of Parental Severity of Discipline.* "How often do your parents punish you?" (Negative item.)
12. *Lack of Parental Overprotection.* "Do your parents try to protect you too much against difficulties or dangers?" (Negative item.)
13. *Lack of Parental Overinsistence on Achievement.* "Are your parents always after you to work hard to become a success?" (Negative item.)
14. *Sympathetic Encouragement of Achievement.* "Do your parents inspire you to want to develop your abilities?"
15. *Regularity of Home Routine.* "Are your meals served at regular hours?"
16. *Harmony of Parents.* "Do your parents openly show affection or consideration for each other?"

The students were also asked to provide sociological data on parents' marital status, occupation, and education and on the ages and sexes of their siblings. In addition, each student was asked to respond to three checklists: (a) a list of activities (e.g., "music," "insects or beach life," "fish or hunt," "billiards") to be checked as to which ones his parents had taught him or developed his interest in, yielding an "intellectual," a "nonintellectual," and a total score; (b) a list of 11 high school goals or values (in the areas

⁶ Unpublished Home Life Scale, developed for use in a study of family morale conducted by Dr. Leland H. Stott. We wish to thank Dr. Stott for making this instrument available to us.

⁷ Unpublished Youth Planning Survey questionnaire, developed by Dr. Fred L. Strodtbeck and his associates for use in a study of achievement motivation and its correlates. We wish to thank Dr. Strodtbeck for making a copy of this questionnaire available to us.

of academic achievement, athletics, social relations and status, general adjustment, morals, and religion) to be checked as to which his parents considered "most" and "least" important (three each to be checked) for him to attain; (c) a similar list of 11 adult goals (e.g., "be outstanding in your occupation," "live a happy life," etc.) to be checked the same way.

The questionnaire also included four open-ended questions: "What sort of person is your mother?" "What sort of person is your father?" "What do you like most about your home and family?" "What would you like to change about your home and family?"

The following measures of hypothesized mediating variables (between family relations and school achievement) were obtained: (a) a six-item attitude scale (with a five-point scale for each item) designed to measure Negative Attitudes to Teachers,⁸ which yielded a corrected odd-even reliability of .69; (b) a similar six-item attitude scale designed to measure Negative Attitudes to School,⁸ which yielded a corrected odd-even reliability of .85; (c) a 72-item interest-inventory scale (with dichotomous items) designed to measure Interest in Intellectual Activities⁹ (esthetic activities, social problems, natural science, and formal symbol manipulation such as mathematics and puzzles), which yielded a corrected odd-even reliability of .77.

Analysis of Data

The internal consistency reliability of each Family Relations Scale and of the Over-all Family Morale Scale was evaluated by computing an odd-even product-moment reliability coefficient. Differences between high- and under-achievers on each scale and on the Over-all scale were tested for significance by a median test.

Group differences in checklist responses were tested for significance by a median test for scale scores and by a chi square test for single items.

As for the open-ended questions,¹⁰ response categories potentially differentiating between the two achievement groups were determined by studying an inspection sample. These categories were cross-validated by being scored¹¹ "blind" on another sample of 66¹² boys in each group with groups equated for grade in school, socioeconomic status, and intelligence. Differences between groups were tested for significance by chi square and median tests.

Pearson product-moment correlations were obtained between the Over-all Family Morale scale and the several measures of mediating variables

⁸ Adapted from the Brown-Holtzman Survey of Study Habits and Attitudes (S.S.H.A.), published by the Psychological Corporation. We are grateful to the Psychological Corporation for permission to make use of the S.S.H.A. for this purpose.

⁹ Adapted from an inventory developed by J. P. Guilford and his associates (6, 7). We are grateful to Dr. Guilford for making the inventory available to us for this purpose.

¹⁰ The analysis of the open-ended questions in the student questionnaire was conducted by J. Robert Wallace under Morrow's supervision.

¹¹ In view of the low validity obtained in this cross-validation, time was not invested in having a second judge score the material to provide a measure of interjudge agreement.

¹² This sample included 30 additional 1956-1957 freshmen in each group.

CHILD DEVELOPMENT

concerned with attitudes toward teachers, school, and intellectual activities. These correlations were computed separately for the total sample, for high-achievers only, and for under-achievers only. In addition, differences between achievement groups on these mediating variables were tested for significance by a median test.

All tests of significance used were two-tailed.

RESULTS

Family Relations Scales

Results on the Family Relations Scales are presented in Table 1. The internal consistency reliability of the 16 six-item scales appeared to be sufficiently satisfactory for purposes of group comparison: all but three of the corrected odd-even coefficients were above .70. The 90-item Over-all Family Morale Scale yielded a reliability of .97.

TABLE 1
FAMILY RELATIONS SCALES: MEDIAN TESTS*

Scale Title	r†	PER CENT ABOVE MEDIAN		p
		Highs (N=48)	Lows (N=48)	
Family Sharing of Recreation76	69	44	.02
Family Sharing of Confidence & Ideas84	63	35	.01
Family Sharing in Making Decisions88	60	44	ns
Parental Approval56	73	33	.001
Parental Affection88	60	42	ns
Parental Trust73	60	25	.001
Parental Approval of Peer Activities94	71	42	.01
Student Acceptance of Parental Standards69	52	25	.01
Student Affection & Respect toward Parents ..	.91	58	44	ns
Lack of Parental Overrestrictiveness63	56	29	.01
Lack of Parental Severity of Discipline70	69	42	.01
Lack of Parental Overprotection77	52	56	ns
Lack of Parental Overinsistence on Achievement	.75	63	46	ns
Parental Encouragement of Achievement74	60	40	.05
Harmony of Parents (N=40)72	63	48	ns
Regularity of Home Routine76	52	46	ns
Over-all Family Morale97	67	33	.001

* Two-tailed tests.

† Odd-even reliability coefficient, corrected by Spearman-Brown formula.

Nine of the 16 six-item scales differentiated between the high- and under-achieving groups in the predicted direction beyond the .05 level, seven of these nine beyond the .01 level. The Over-all Family Morale Scale differentiated beyond the .001 level.

High-achievers more often than under-achievers (a) described their families as typically sharing recreation, ideas, and confidences; (b) described their parents as approving and trusting (the areas of sharpest difference between the two groups), affectionate, encouraging (but not pressuring) with respect to achievement, and relatively nonrestrictive and nonsevere; and (c) described themselves as accepting their parents' standards.

An equal majority of both groups described their parents as having a relatively harmonious relationship, portrayed their homes as having a fairly regular routine, denied that they were either seriously overprotected or excessively pressured to achieve, and said they felt considerable respect and affection for their parents.

Even in the areas of greatest difference, however, there was considerable overlap between the two groups. On the Over-all Family Morale Scale nearly a third of the high-achievers scored below the median, and nearly a third of the under-achievers scored above the median.

Sociological Data

The two groups did not differ significantly in any of the sociological factors on which data was obtained. The data contradict the stereotyped notion that a mother's working outside the home inevitably leads to neglect which conduces to poor school performance (and other dire consequences). Actually 47 per cent of the high-achievers (as against 37 per cent of the under-achievers) reported that their mothers were working outside the home!

Check Lists

The two groups showed essentially no significant differences in their responses to the check lists.

Open-Ended Questions

In the inspection sample 34 potentially differentiating response categories were noted, of which 25 were grouped in five "high" combined categories and nine were grouped in three "low" combined categories. In the cross-validation only two "high" combined categories (out of eight combined categories) and four single "high" categories differentiated significantly (in the predicted direction).

Fifty-one per cent of the high-achievers as against 33 per cent of the under-achievers ($p < .05$) were scored in the combined category, "Positive references to intrafamily relationships." Two single categories in this grouping also differentiated significantly: "References to parental interest in family or student" (29 per cent to 12 per cent, $p < .01$); and "References to parents' outgoing, positive shaping of student's development" (12 per cent to 0 per cent, $p < .02$).

The combined category, "References to 'Golden Rule' virtues of parents" was scored significantly more often for high-achievers (56 per cent to 29 per

CHILD DEVELOPMENT

cent, $p < .01$). Two single categories in this grouping also differentiated significantly: "Parents described as 'considerate' or 'thoughtful'" (17 per cent to 4 per cent, $p < .05$); and "Parents described as 'understanding'" (36 per cent to 17 per cent, $p < .02$).

These results are consistent with those on the Family Relations Scales.

Mediating Variables

The correlation between Over-all Family Morale and Negative Attitudes to Teachers was $-.67$ for the total sample, $-.46$ for high-achievers only, and $-.60$ for under-achievers only. The corresponding correlations between Family Morale and Negative Attitudes to School were $-.69$, $-.32$, and $-.65$. The correlations between Family Morale and Intellectual Interests were $.57$, $.31$, and $.49$. All of the obtained correlations are significant at the $.05$ level or better. On each of the three measures of mediating variables the two achievement groups differed at the $.001$ level in the expected direction.

These results are consistent with the hypothesis that supportive family relations foster academic achievement *via* promoting positive attitudes toward teachers, school, and intellectual activities, as symbols of the adult world of parents. However, other interpretations of the direction of causality are not excluded.

SUMMARY

The reported family relations of 48 high school boys of superior intelligence making high grades were compared with those of a group making mediocre or poor grades, equated for grade in school, socioeconomic status, and intelligence. The main measuring instrument was a set of 16 six-item questionnaire scales (with a four-point scale for each item), on which each subject was asked to describe his family relations. The students also provided sociological data, check-list data on parental goals for the student, and open-ended question data on conceptions of parents. In addition, attitude scale data were obtained on variables hypothesized to mediate the influence of family morale on student achievement.

The results supported the hypotheses that bright high-achievers' parents reportedly engage in more sharing of activities, ideas, and confidences; are more approving and trusting, affectionate, and encouraging (but not pressuring) with respect to achievement; are less restrictive and severe; and enjoy more acceptance of parental standards by their youngsters. Not supported were hypotheses that under-achievers' families show more over-protectiveness, more high-pressure for achievement, more parental disharmony, more irregularity of home routine; differences in goals for their youngsters; or differences in sociological factors such as parents' marital status, current occupation of either parent, or number and ages of siblings.

The results also supported the hypothesis that family morale fosters academic achievement among bright high school boys *via* fostering positive

attitudes toward teachers and toward school and interest in intellectual activities, as mediating variables. However, other hypotheses as to the direction of causality are not ruled out.

REFERENCES

1. BLOCK, V. L. Conflicts of adolescents with their mothers. *J. abnorm. soc. Psychol.*, 1937, 32, 193-206.
2. COLEMAN, H. A. The relationship of socio-economic status to the performance of junior high school students. *J. exp. Educ.*, 1940, 9, 61-63.
3. COLLINS, J. H., & DOUGLAS, H. R. The socio-economic status of the home as a factor in success in the junior high school. *Elem. Sch. J.*, 1937, 38, 107-113.
4. CONKLIN, A. M. Failures of highly intelligent pupils. *Teach. Coll. Contrib. Educ.*, 1940, No. 792.
5. GOUGH, H. G. The relationship of socio-economic status to personality inventory and achievement test scores. *J. educ. Psychol.*, 1946, 37, 527-540.
6. GUILFORD, J. P., CHRISTENSEN, P. R., BOND, N. A., JR., & SUTTON, M. A. *Technical appendices to the research bulletin, A factor-analytic study of human interests*. Beverly Hills, Calif.: Sheridan Supply Co., 1953.
7. GUILFORD, J. P., CHRISTENSEN, P. R., BOND, N. A., JR., & SUTTON, M. A. A factor analysis study of human interests. *Psychol. Monogr.*, 1954, 68, No. 4.
8. HARRIS, D. Factors affecting college grades: a review of the literature, 1930-1937. *Psychol. Bull.*, 1940, 37, 125-161.
9. HATTWICK, B. W., & STOWELL, M. The relation of parental over-attentiveness to children's work habits and social adjustments in kindergarten and the first six grades of school. *J. educ. Res.*, 1936, 30, 169-176.
10. HAVIGHURST, R. J., & TABA, H. *Adolescent character and development*. Wiley, 1949.
11. ITKIN, W. Some relationships between intra-family attitudes and pre-parental attitudes toward children. *J. genet. Psychol.*, 1952, 80, 221-252.
12. JONES, E. S. The probation student: what he is like and what can be done about it. *J. educ. Res.*, 1955, 49, 93-102.
13. KIMBALL, B. The sentence-completion technique in a study of scholastic achievement. *J. consult. Psychol.*, 1952, 16, 353-358.
14. KIMBALL, B. Case studies in educational failure during adolescence. *Amer. J. Orthopsychiat.*, 1953, 23, 406-415.

CHILD DEVELOPMENT

15. LEIBMAN, O. B. The relationship of personal and social adjustment to academic achievement in the elementary school. Unpublished doctoral dissertation, Columbia Univer., 1953.
16. LEWIS, W. D. A comparative study of the personalities, interests, and home backgrounds of gifted children of superior and inferior educational achievement. *J. genet. Psychol.*, 1941, 59, 207-218.
17. MUSSELMAN, J. W. Factors associated with the achievement of high school pupils of superior intelligence. *J. exp. Educ.*, 1942, 11, 53-68.
18. RICKARD, G. The relationship between parental behavior and children's achievement behavior. Unpublished doctoral dissertation, Harvard Univer., 1954.
19. TIBBETS, J. R. The role of parent-child relationships in the achievement of high school pupils. *Dissert. Abstr.*, 1955, 15, 232.
20. WALSH, A. *Self-concepts of bright boys with learning difficulties*. Columbia Univer., 1956.

PEER POPULARITY AND DEPENDENCE ON ADULTS IN PRESCHOOL-AGE SOCIALIZATION

BOYD R. McCANDLESS,*

State University of Iowa

CAROLYN B. BILOUS and HANNAH LOU BENNETT

University of Hawaii

The present research was designed to investigate the interaction between popularity, dependency on adults, and sex in preschool-age children. Marshall and McCandless, studying 38 preschool children in two groups of the Iowa Child Welfare Research Station's Laboratory Preschools, found that "dependence on adults in the preschool situation accompanies relatively low social status and participation" (2, p. 419). In that investigation dependency was defined as simply the total number of interactions with adults during a prescribed period of observation. In addition to this over-all negative relation between popularity and dependence, Marshall and McCandless found that sex differences significantly characterized the interaction between popularity and dependence; girls showed a median correlation between popularity and dependence of about $-.70$, while the median correlation for boys was about $-.40$.

The present study, then, began as an investigation of why (and if) dependency interferes more with popularity for girls than boys.

The correlations reported by Marshall and McCandless imply that a relatively high percentage of adult contacts involve some sort of socially inap-

* Iowa Child Welfare Research Station, State University of Iowa, Iowa City; on leave, University of the Punjab, Lahore, West Pakistan.

appropriate dependency and that the girl-adult contacts are more characterized by this than the boy-adult contacts. From observation of preschool play it seemed likely that one frequent type of socially inappropriate dependency contact with adults is seeking adult intervention to solve peer-conflict situations. It was hypothesized that seeking adult intervention in peer-conflict situations would interfere with popularity and that a higher percentage of the adult contacts would be comprised of attempts to seek adult intervention in peer-conflicts for girls than for boys. This would in part account for the differential tendency for dependency to interfere more with popularity for girls than for boys.

Heathers (1) has suggested that the total number of contacts with adults can be usefully classified into two separate categories: (a) emotional dependency and (b) instrumental dependency. Such a classification seemed useful for the present investigation. It is here suggested that children who are emotionally dependent upon adults are motivated to learn social skills necessary to gain emotional satisfactions from adults, rather than to learn skills necessary to gain emotional satisfactions from peers. Thus, a differential tendency for girls more than boys to be *emotionally* dependent upon adults may also account for the previous findings that dependency interferes more with popularity for girls than for boys. Instrumental dependency, on the other hand, since it is unrelated to interpersonal, emotional need-satisfaction, should not interfere with popularity.

It would also be expected that the differential responses of girls and boys to mother-figures (teachers) had been learned in the home. Thus, it was expected that the mothers of girls would more frequently act as problem solvers in conflict situations and be more nurturant and supportive of the child in conflict situations than would the mothers of boys.

OBTAINED MEASURES

Popularity

Teacher judgments. The head teacher and her three student assistant teachers of each of the three preschool groups involved in this study were asked to nominate the four best friends of each child in the respective groups. Thus, a total of 16 nominations for each child were used. The index of popularity used was the simple count of number of mentions as "one of the four best friends" of other children. This was then converted to a ratio of possible mentions. Average interteacher correlation was .54.

Free play observations. As an alternative (and probably sounder) index of popularity, the time spent during observation by each child in associative and cooperative play with every other child was tabulated. The best friends of each child were considered to be the four with whom he spent the most time in such contact. That child who was one of the four best friends of the most children was designated as most popular; the one who least fre-

quently was one of the four best friends of others was designated least popular. Thus, popularity was defined as "being sought after" by peers.

Child Behavior with Adults (Teachers)

Fifteen five-minute samples of data for each child were obtained by the first author according to a semiformal time sample observational technique developed by him during observations conducted prior to the beginning of the present study. Observational focus was on peer conflicts and their solution, although notations of a number of social interactions of the child during each five-minute period were also made. Reliability of observations between the first and third authors for approximately four hours of observation was 92 per cent exact agreement of categorization. As a further check on the technique, percentage of agreement between the first author and a psychology graduate student who had no prior experience with observation of children was 65 during an initial two-hour observation and, after discussion, rose to 93 for a second and subsequent two-hour observational period.¹

For the present study, the following observed events were tabulated for each child: (a) number of contacts made with adults, (b) total time spent interacting with adults, (c) number of times instrumental "nurturance" was solicited from adults, (d) number of times emotional "nurturance" was solicited from adults, and (e) the number of times that a child requested or threatened to request help from the teacher.

Mother Behavior with Child

Since it was not feasible actually to obtain observational data concerning mother-child interactions, reliance was placed upon situation-bound maternal interviews. Mothers were asked to describe the last two instances of conflict in which they had observed their children and to tell what they were about, who was involved, what had gone on, and what their role had been. If possible, one of these conflicts was between the child and his siblings, while the other conflict involved a neighborhood playmate or playmates. For only children, or children who had only substantially older siblings, two neighborhood conflicts were elicited. The second and third authors conducted all interviews and secured excellent cooperation from the mothers. From typescripts of the interviews 29 categories of maternal response to child conflict were tabulated. The three authors classified 10 interviews with full discussion of their disagreements, then independently coded the remaining interviews. Percentages of exact agreement on categorization were: between authors 1 and 2, 81.4; 1 and 3, 70.6; and 2 and 3, 75.5.

Two clinical type scales were then devised to lump the categories for purposes of the present investigation: (a) "*Let Them Fight Their Own*

¹ The authors appreciate the work of Mr. Mark Beals in establishing reliability of observations.

CHILD DEVELOPMENT

Battles." This scale lumps together those maternal responses where there is no active maternal intervention. Sample behaviors are: encourages child to defend self, ignores the issue or situation, encourages child to work through own conflicts. (b) "*Active Intervention with Constructive Intent.*" Sample items are: suggests compromise with explanation and/or reasoning, suggests withdrawal from conflict for ethical reasons, encourages child to come to mother.

SUBJECTS

The subjects (Ss) were 26 Caucasian youngsters, referred to in Hawaii as *haoles*, and 36 *nonhaole* children. The latter were of predominantly Oriental ethnic origin. Two children—one a blind girl, the second a mentally retarded boy—were dropped from the study, leaving a final *N* of 60, 28 boys and 32 girls, of whom 26 were *haole*. Ages ranged from 3 years, 6 months to 5 years, 3 months, the mean being 4 years, 5 months. All children attended the University of Hawaii's preschools,² and all but one had been in those preschools for at least six months. This child was the senior author's daughter; she had previously attended preschool on the mainland, and data for her fell close enough to the median of the group for the obtained measures that it was considered legitimate to retain her as an *S*, even though her father was the primary observer for the study.

Maternal interviews were not obtained for three of the children, so that the mothers of only 57 of the children provided information about maternal responses to child conflict.

RESULTS AND DISCUSSION³

The central problem in this study was to explore sex differences and their relations to adult-dependency behavior so as to replicate and throw additional light on a finding originally reported by Marshall and McCandless (2). Marshall and McCandless found dependency, defined grossly as total number of contacts with adults, negatively related to popularity with peers (defined by free play contacts). This negative relation was significantly stronger for girls than boys.

The present study lends confidence to this result, although the relations between gross number of adult contacts and popularity are weaker than Marshall and McCandless obtained for an Iowa population. For both sexes combined in the present Hawaiian sample, r was $-.30$ ($p \leq .05$) when the

² The authors appreciate the courtesy and cooperation of Mrs. Elizabeth Crooker and Miss Mary Rushit, head teachers of the intermediate and older preschool groups, in making it possible to work with their groups.

³ Dr. Charles Truax has been extremely helpful in suggesting methods of data analysis, carrying out statistical procedures, and giving helpful critical attention to the manuscript.

criterion of popularity was free play association and $-.16$ (ns) when teacher judgment was used to define popularity. The correlation (using free play popularity) was higher for girls than boys ($-.38$, $p \leq .05$ for girls; $-.18$, ns , for boys). While this result is in the direction of the Marshall-McCandless findings, the difference between the correlations in the present study is not significant.

In the present study dependency was broken down more refinedly than it had been by Marshall and McCandless. Dependency behaviors were separated, in accord with suggestions by Heathers (1), into *emotional* and *instrumental* dependency. Emotional dependency consists principally of seeking for comfort, affection, and support; instrumental dependency of asking objective help ("Help me put up the ladder"). However, requests for teachers to intervene in conflicts were classified by the present authors as emotional dependency. After considerable observation of behavior, the present authors agreed that the emotional, support-seeking component ("Teacher! Teacher!" or—while crying—"He hit me!") of this behavior was more significant than its instrumental aspect.

The authors hypothesized that emotional dependency accompanies (or precedes) inadequacy with peers. Excessive emotional dependency in a child may mean that he is more focused on adults than children, hence does not develop effective techniques for relating with peers. It may mean that dependency relations with parents are not yet resolved and that, accompanying this, emotionally dependent children are still sufficiently insecure that they fail to relate successfully to other children.

The hypothesis that emotional dependency would interfere with popularity was supported. When the free play criterion of popularity was used, emotional dependency and popularity correlated $-.33$ ($p \leq .05$); teacher-judged popularity and emotional dependency correlated $-.27$ ($p \leq .05$). It can be seen from Table 1 that girls show more total emotional dependency than boys (although the significance of this difference disappears when "asking teacher intervention" is eliminated from the category). However, rather surprisingly, "asking teacher intervention" did not interfere with popularity. The authors had expected that the child who indulged extensively in such behavior would be regarded by his peers as a "ratter" or that such behavior would interfere with the normal, competent solving of conflicts that they believed accompanied popularity. But the correlation between amount of "asking intervention" and popularity (free play) was $-.08$ (ns); with popularity (teacher judgment) it was $.04$.

Thus, no clear answer is provided for the question of why emotional dependency tends to interfere more with popularity for girls than boys: girls are not significantly different from boys in amount of emotional dependency when "asking intervention" is eliminated from the category; they *do* request much more intervention from teachers than boys (r_{bis} is $+.58$ for "asking intervention" and sex, with girls at the positive pole); yet asking

CHILD DEVELOPMENT

TABLE I

SEX DIFFERENCES IN BEHAVIOR AS DETERMINED BY OBSERVATION*

	MEDIAN FREQUENCY OF OCCURRENCE†		χ^2	<i>p</i>
	Boys	Girls		
Asks teacher to solve conflict	0.6	2.2	20.17	.01
Total dependency	3.0	4.3	1.61	<i>ns</i>
Emotional dependency	1.6	3.8	5.34	.05
Instrumental dependency	1.4	0.9	0.31	<i>ns</i>
Conflicts engaged in	11.5	9.4	2.41	<i>ns</i>
Conflicts initiated	6.0	2.8	7.04	.01
Token or no resistance to conflict (%)	33.0	59.0	5.44	.02
Prolonged resistance to conflict (%)	21.7	5.2	2.41	.20
Persists in activity after conflict (%)	98.0	82.0	8.42	.01
Changes activity after conflict (%)	9.2	26.2	7.04	.01
Contacts with other children	51	48	0.34	<i>ns</i>
Contacts with adults	9.2	8.7	0.12	<i>ns</i>
Instrumental nurturance offered by adults	3.0	3.0	0	<i>ns</i>
Emotional nurturance offered by adults	0.8	0.9	0.21	<i>ns</i>
Teacher popularity ranking (weighted score)	50	48	0.26	<i>ns</i>

* The Yates correction for one degree of freedom has been used in computing chi square.

† Legends indicate when percentages or weighted scores were used.

intervention is not related to popularity, while emotional dependency is significantly related in a negative direction, and more strongly so for girls than boys.

The most plausible speculation that occurs to the authors concerns cultural differences in treatment of emotional dependency between girls and boys: it is moderately acceptable for girls, relatively unacceptable for boys. Girls, then, who have difficulty in peer relations are likely to "run to the teacher," but boys with similar problems (who have received discouragement for emotional dependency) seek other methods of compensating, perhaps in some cases by ruggedly independent or aggressive behavior, so that what relation there might be for them between emotional dependency and popularity is attenuated and possibly in some cases negated (i.e., the culture forces "compensation for emotional dependency" on boys, but not on girls).

It was predicted that boys and girls would not differ in instrumental or total dependency. Table I shows that this prediction was confirmed. Neither was it expected that instrumental dependency would be related to popularity. This proved to be the case: free play popularity and instrumental dependency correlated $-.14$; popularity as judged by teachers and instrumental dependency correlated $-.08$.

Another hypothesis was that, in addition to finding that "asking intervention" from teachers would be more frequent for girls than boys, the maternal interviews would show that girls had learned such behavior from their mothers' methods of handling their conflicts: i.e., mothers would let boys "go it alone" more frequently, while they would set themselves up more often as problem solvers for their daughters. But biserual correlations with these two maternal "scales" were only $-.09$ and $-.15$, respectively. This lack of relation is most likely due to the low reliability of categorizing the interview, lack of validity in the interview, the limited sampling of mother-child interactions, or a combination of these factors.

Table 1 supplies certain other data of interest, although bearing less directly on the major purposes of the study: Boys, significantly more often than girls, initiate conflicts. That boys are more task-oriented than girls is indicated by the greater frequency with which they persisted in the activity in which they were engaged when conflict occurred and the lesser frequency with which they changed to another activity following conflict. Girls more often than boys offer little or no resistance to conflict. The sexes do not differ in popularity as judged either by teachers or from free play, nor in either instrumental or emotional nurturance offered by adults, nor in number of contacts with adults (a rough index of total dependency). It is the *nature* of these adult contacts that seems to make the difference.

SUMMARY

The central problem of this study was to investigate "refinements" of dependency behavior as they vary by sex and affect children's relations with their peers. The authors were also interested in the relation between "general" adult-dependency and popularity. A general postulation was that girls would show more emotional dependency than boys and that emotional—but not instrumental—dependency on adults would interfere with social relations with peers.

Twenty-six Caucasian (*haole*) and 34 non-Caucasian (*nonhaole*) University of Hawaii preschool children were used as subjects; 23 were boys, 32 girls.

Popularity and emotional dependency were found to be negatively related, as were "total adult contacts" and popularity. This over-all index of adult dependency interfered more with girls' than with boys' popularity, although the difference between sexes was not as great as had been found in an earlier study done with University of Iowa preschoolers.

Other findings indicated that mothers did not intervene more frequently in daughters' than in sons' conflicts, although girls more frequently than boys asked teachers to solve their conflicts. Girls showed more total emotional dependency than boys, although the difference became nonsignificant when "asks teacher intervention" was eliminated from the category of emotional dependence. Girls were less likely to resist conflicts and were

CHILD DEVELOPMENT

more likely to change activity following conflict than boys. Girls initiated fewer conflicts, but engaged in almost as many as boys. There were no sex differences in instrumental or total dependency, and instrumental dependency did not interfere with popularity. There were no sex differences in total or in instrumental dependency or in popularity, whether this was judged by teachers or measured by associations during free play.

REFERENCES

1. HEATHERS, G. Emotional dependence and independence in nursery school play. *J. genet. Psychol.*, 1955, 87, 37-57.
2. MARSHALL, H. R., & McCANDLESS, B. R. Relationships between dependence on adults and social acceptance by peers. *Child Develpm.*, 1957, 28, 413-419.

SEX OF EXPERIMENTER AND SEX-TYPED BEHAVIOR OF YOUNG CHILDREN

L. J. BORSTELMANN*¹

Stanford University

Although in recent years considerable attention has been given to the study of parents as significant contributors to child behavior patterns and tendencies (3, 8, 9, 12, 13, 15, 16), there has been little evident awareness of experimenters as adult interactors varying in attributes or behaviors that may affect responses elicited from children. Whereas variations in parental attitudes and behaviors are often hypothesized as critical antecedents, experimenters seem to be regarded as generalized adult figures, impersonal behavior initiators, and data recording machines. If the child is indeed responsive to such qualities in his parents as warmth, permissiveness or restriction, sex roles, etc., then he should build up certain expectations and anticipatory sets associated with the occurrence of such qualities in other significant adults. Yet studies designed to manipulate child behavior have generally failed to consider or control for relevant experimenter characteristics. Some attention has been directed to the presence or absence of adults as a factor affecting child behavior (14), but the effects of *E* attributes or behaviors have been largely neglected.

One of the more obvious, readily defined *E* attributes that might be considered to influence a child's responses is sex status. Interactions with male versus female experimenters may be expected to involve some differences in response sets of the experimental children. A few recent studies have varied sex of *E* as part of the experimental design. Gewirtz and his colleagues (6) in two studies of attention seeking with variation of adult

* Department of Psychology, Duke University, Durham, North Carolina.

¹ This study was conducted while the author was a Research Fellow of the United States Public Health Service. Appreciation for advice, guidance, and cooperation is acknowledged to Robert R. Sears, Director, Laboratory of Human Development, and Edith Dowley, Director, Nursery School, Stanford University, and their staffs.

availability found a significant cross-sex pattern of child response, both boys and girls making more dependency bids to the opposite sex *E*. Rosenblith (11) found a male *E* to be more effective than a female *E* in eliciting imitation learning. Child (2) found a greater frequency of distant versus immediate goal choices with a female *E*. All of these investigators employed only one *E* of each sex and are properly cautious in noting that the obtained effects might be due to *E* characteristics other than sex difference. Such equivocal conclusions represent the usual state of affairs when the issue is peripheral to the primary interests of the investigators. We are left either with an unwarranted generalization from an *N* of 1 or with the tantalizing suggestions of an important effect that is unfortunately not verified by representative sampling of the presumably relevant variables. Hammond (7), following Brunswik's notion of representative versus systematic design, has pointed out that unwarranted conclusions concerning the effects of examiners occur as a result of failure to establish the crucial independent variable. The present study was designed to implement this *E* sampling requirement with respect to the effects of sex of *E* upon the sex-typed behavior of young children.

In considering the possible effects of *E* sex status upon the degree of the child's appropriately sex-typed responses, two contradictory hypotheses are suggested: (a) that an *E* of the same sex as the child will enhance sex-appropriate responses of the child by providing reinforcement as an identification model; or (b) that an *E* of the opposite sex to the child will enhance sex-appropriate responses by providing reinforcement for assertion of the complementary role. The possibility that either or both hypothesized effects might operate within a given group of children led to our investigation of sex of *E* as a variable in a test-retest design. The alternative possibilities of *E* influence will be evident as follows:

1. If the male versus female status of *Es* produces a constant effect among boys or girls, this will be reflected in different degrees of sex-appropriate behavior according to sex of *E*.
2. If sex of *E* produces a variable effect among children, then test-retest agreement will be greater when sex of *E* is unchanged than when sex of *E* is reversed.

PROCEDURE AND SUBJECTS

The study design involved initial and repeat administration of three experimental measures of sex-typed behavior under conditions of *E* sex constancy and of *E* sex reversal. Each of eight male and eight female *Es* administered all three measures to two boys and two girls. All children had a change of *E* for readministration, but half of the subjects of each sex had a new *E* of the same sex as the initial *E* and half had a shift in sex of *E*. Thus, the 32 boys and 32 girls were assigned randomly to four sequences of *Es*: male-male, male-female, female-female, and female-male.

The experimental tasks were three measures of sex-typed behavior used by Sears and his group in their current study of identification in young children: the It Scale developed by Brown (1) and two instruments especially constructed for the project, Toy Preference (TP) and the Pictures Test (Pix). TP was adapted from a procedure of Rabban (10), using seven male and seven female toys with the child instructed to make eight consecutive choices of preferred items. This technique has proven to have consistently high interest value for nursery school children. The Pictures Test, adapted from a method used by Fauls and Smith (5), is a set of 12 items with separate forms for each sex, wherein the child is depicted in combinations of male or female, adult or child activities in the presence of both parent figures. The subject is asked to make a simple choice of preferred activity in a series of paired comparisons. This technique provides a measure of adult role preference as well as sex role preference.

The 64 child *Ss* were all enrolled in the University Nursery School, where the study was conducted. They ranged in age from 3 years, 4 months to 5 years, 0 months with a median age of 4-2 and no age difference between boys and girls. The 16 *Es* were university students in child study courses, familiar with the nursery school. All *Es* were pretrained in the use of the experimental measures.

RESULTS

The first hypothesis regarding any general effects of sex of *E* upon degree of children's sex-typed responses was examined in terms of differences between male and female *Es* on initial administration of the three measures and differences between the two readministration conditions of constancy versus change in sex of *E*. There were no significant mean differences due to sex of *E* for boys or girls on the initial measures. There were no significant mean differences between test 1 and test 2 for boys or girls regardless of *E* sequence. Thus, there is no evidence that sex of *E* produces any constant effect of either increasing or decreasing the sex-appropriate responses of boys or girls. Boys had a consistently and significantly higher mean sex-appropriate score on the Toy Preference and It Scale, but not on the Pictures Test.²

The second hypothesis of variable effects of sex of *E* upon child response was examined in terms of any differences between retest reliabilities for the experimental conditions of constancy versus reversal of sex of *E*. If there is interaction between sex of *E* and identity of *S* (as distinct from that between sex of *E* and sex of *S*), then the reliabilities will be lower for male-female and female-male sequences than for male-male and female-female. As may be seen from the results as presented in Table 1, there are no significant differences among the Pearson correlations for any of the measures.

² Information about the sex-typed measures, scoring criteria, and distributions will be provided upon request to the Laboratory of Human Development, Stanford University.

CHILD DEVELOPMENT

TABLE 1

TEST-RETEST RELIABILITY COEFFICIENTS FOR THREE SEX-TYPING
MEASURES WITH STABILITY AND REVERSAL
OF SEX OF EXPERIMENTER (*E*)

Experimental Condition	N	TP		IT		PIX	
		B	G	B	G	B	G
<i>E</i> Sex Constant	16	.60 [§]	.56 [§]	.79 [†]	.63 [‡]	.38	.19
<i>E</i> Sex Reversed	16	.73 [†]	.74 [†]	.81 [†]	.66 [‡]	.38	.52 [§]
Total	32	.76*	.62*	.80*	.64*	.37 [§]	.38 [§]

* $p < .001$.† $p < .01$.‡ $p < .02$.§ $p < .05$.

Again there is no evidence that sex of *E* has any significant effect on the sex-typed choices of the subjects. The reliabilities actually tend to be slightly higher with *shift* in sex of *E* for both sexes on Toy Preference and for girls on the Pictures Test.

Two further aspects of the data were examined: age effects and agreement among measures. Since age of *Ss* is distributed randomly among *E* sequences, any age differences in sex-appropriate responses may be ascertained. For this purpose both boys and girls were separated into older and younger groups in accordance with the median split of 4 years, 2 months. Older boys showed a general trend of more sex-appropriate responses, but the difference was significant only for the It Scale. There were no age differences in retest consistency for boys. Girls reflected consistent age differences only on the It Scale, with older girls more sex-typed and younger girls less reliable.

Finally, the intercorrelations among the three measures as shown in Table 2 are all positive but not consistently significant. TP and Pix are significantly correlated for both sexes on initial measures, but the significance

TABLE 2

INTERCORRELATIONS AMONG THE SEX-TYPING MEASURES

	Boys (<i>N</i> = 32)		Girls (<i>N</i> = 32)	
	Test 1	Test 2	Test 1	Test 2
TP - It33	.35*	.25	.30
TP - Pix44*	.33	.51 [†]	.09
It - Pix26	.40*	.28	.39*

* $p < .05$.† $p < .01$.

drops out on the retest measures. It and Pix are not significantly related on test 1 but become so on test 2 for both boys and girls.

DISCUSSION

Male or female status as an attribute of *Es* has been investigated for possible effects upon sex-appropriate behavior of young boys and girls with negative results. Thus, the sex role aspects of the adult-child interaction must be considered negligible with respect to the experimental measure of sex-typed behaviors. This conclusion does not suggest that Gewirtz' and Rosenblith's findings of differences related to sex of *E* are fortuitous, since they were studying behaviors of dependency and imitation that were demonstrably a function of *E* behavior. In the present study the tasks employed were intrinsically interesting to the children, and the *Es* were permissive with no cues to reinforce certain responses. Thus, the situation was structured to maximize the child's recourse to his own personal preferences. There were no known external referents that might direct responses, other than sex of *E*, and this was apparently irrelevant to the children. There of course remains the possibility that some differential effects due to *Es* did occur, but we have sampled only the dimension of sex status and our sample size does not permit analysis of individual *E* effects.

The negative results of the present study do underscore the necessity of adequate sampling of *Es* for determination of any experimental effects attributed to *E* differences. As Gewirtz and Rosenblith both note, their differences between one male and one female *E* may well be due to many individual characteristics other than their obvious sex status. In the present instance we have systematically studied sex of *E* as a relevant experimental variable and have found no effects evident. We do not conclude therefrom that children are indifferent to sex of adults but rather that any inferences about the effects of *E* attributes or behaviors upon child response can only be justified on the basis of an adequate sampling of *Es* on the relevant characteristics with reference to given behavior systems of the child.

The present findings confirm the reliability estimates reported by Brown for the It Scale and by Rabban for a similar toy preference technique. The low reliabilities of the Pictures Test are in line with other uses of paired comparisons procedures with nursery school children, as reported by Clifford (4).

Since the intercorrelations among the three measures employed are quite low, there remains some question about the generality of a concept of sex-typing at the nursery school age, at least insofar as toy and play preferences are concerned.

SUMMARY

Alternative hypotheses of constant or inconsistent effects of experimenter sex difference upon the responses of young children to three measures of

CHILD DEVELOPMENT

sex-typing were studied by means of a test-retest design under two conditions of change or constancy of sex of experimenters. The results yielded clear evidence that sex of experimenter was not a significant variable. Sex and age differences in degree of appropriate sex-typing, and the limited communality among the measures, were noted.

REFERENCES

1. BROWN, D. G. Sex-role preference in young children. *Psychol. Monogr.*, 1956, 70, No. 14.
2. CHILD, I. L. Children's preference for goals easy or difficult to obtain. *Psychol. Monogr.*, 1946, 60, No. 4.
3. CHILD, I. L. Socialization. In G. Lindzey (Ed.), *Handbook of social psychology*. Addison-Wesley, 1954. Pp. 655-692.
4. CLIFFORD, E. Ordering of phenomenon in a paired comparisons procedure. *Child Developm.*, 1959, 30, 381-388.
5. FAULS, L. B., & SMITH, W. D. Sex-role learning of five-year-olds. *J. genet. Psychol.*, 1956, 89, 105-119.
6. GEWIRTZ, J. L., BAER, D. M., & ROTH, C. H. A note on the similar effects of low social availability of an adult and brief social deprivation on young children's behavior. *Child Developm.*, 1958, 29, 149-152.
7. HAMMOND, K. R. Representative vs. systematic design in clinical psychology. *Psychol. Bull.*, 1954, 51, 150-159.
8. LEVY, D. M. *Maternal overprotection*. Columbia Univer., 1943.
9. MILLER, E. R., & SWANSON, G. E. *Inner conflict and defense*. Holt, 1960.
10. RABBAN, M. Sex-role identification in young children in two diverse social groups. *Gent. Psychol. Monogr.*, 1950, 42, 81-158.
11. ROSENBLITH, J. Learning by imitation in kindergarten children. *Child Developm.*, 1959, 30, 69-80.
12. SEARS, R. R., MACCOBY, E. E., & LEVIN, H. *Patterns of child rearing*. Row, Peterson, 1957.
13. SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, P. S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-236.
14. SIEGEL, A., & KOHN, L. Permissiveness, permission and aggression: the effect of adult presence or absence on aggression in children's play. *Child Developm.*, 1959, 30, 119-130.
15. WATSON, R. I. *Psychology of the child*. Wiley, 1959.
16. WHITING, J. W. M., & CHILD, I. L. *Child training and personality*. Yale Univer., 1953.

MATERNAL ACCEPTANCE OF RETARDED CHILDREN: A QUESTIONNAIRE STUDY OF ATTITUDES AND RELIGIOUS BACKGROUND

G. H. ZUK,* RALPH L. MILLER, JOHN P. BARTRAM,
and FREDERICK KLING¹

St. Christopher's Hospital for Children, Philadelphia

This study proposes to examine the relation of religious background to maternal acceptance of children. In this case, the children are retarded. Maternal acceptance will be explored through means of a questionnaire tapping attitudes of the mothers toward their children.

Guilt as an important psychological process in parent-child relations has recently been examined in a study by one of the writers (27). Catholic mothers were judged to be more acceptant of their retarded children than non-Catholic mothers. This result was interpreted in the context of the Catholic dictum that parents not blame themselves for bearing a retarded child. Many writers have isolated guilt as a major factor impeding parental acceptance of handicapped children. Sarason (16), Scheimo (19), Waterman (24), Rheingold (12), Hastings (6), and Stone (21) have made this point in reviewing problems in acceptance of retarded children. As it affects the acceptance of cerebral palsied children, it has been mentioned by Perlstein (10), Wishik (25), Moore (9), and Roe (14).

Two recent studies seem to support the finding that parental acceptance of retarded children is greater among Catholics than non-Catholics. One by Saenger (15) reports a significantly higher percentage of institutionalization among Jews than Italians, a predominantly Catholic group. In another, Farber (5) notes that the marital integration of Catholic families is less disturbed by the presence of a retarded child in the home; he concludes that "... non-Catholic families seem more affected than Catholics in a crisis situation" (p. 65).

In their book on the Hutterites, Eaton and Weil (4) point to the high level of acceptance of the retarded in the community. In this sect, the retarded

* St. Christopher's Hospital for Children, 2600 North Lawrence Street, Philadelphia 33, Pennsylvania.

¹ The first three authors are, respectively, psychologist, chaplain, and pediatrician (director) of the Mental Retardation Clinic at St. Christopher's Hospital for Children. The fourth author acted as statistical consultant on the study. The study was supported in part by grants from the following agencies: Children's Bureau, Department of Health, Education, and Welfare; Pennsylvania State Department of Health, Division of Child and Maternal Health; and the Dollfinger-McMahon Foundation. The material presented in this study also served as the basis for an unpublished master's thesis (8).

CHILD DEVELOPMENT

child is not held personally responsible for the mischief he may commit because he is thought not to have reached the age of accountability before God. There seems to be a parallel between the Hutterite and Catholic viewpoints, as described for the Catholics, for example, in a religious tract by Breitenbeck (3).

Some apparently contradictory evidence comes from a recent study by Boles (2). Although his subjects were mothers of cerebral palsied children, this group of children overlaps the retarded greatly in mental ability. Boles analyzed the responses to an extensive personality questionnaire and found that Catholic mothers verbalized more guilt and anxiety and were more socially withdrawn than non-Catholic mothers. This finding applied to Catholic mothers of nonhandicapped children as well, so its interpretation in the light of the handicap variable is ambiguous. However, one may interpret it as evidence for lesser acceptance of their children by Catholic mothers.

This study hopes to provide a step toward clarification of the relation between a cultural factor, religious background, and a psychological concept, maternal acceptance. It hopes to provide evidence bearing on the note, often made in the clinical literature, that guilt impedes parental acceptance of handicapped children. It hopes to provide evidence also bearing on another point made in the clinical literature, that attitudes expressed by parents of handicapped children about these children often may not accurately reflect actual practices.

PROCEDURE AND METHOD

The Questionnaire

A questionnaire was devised to elicit information about (a) the religious practices and (b) attitudes, feelings, and beliefs of mothers of retarded children about retarded children.² Part 1 of the questionnaire was composed of two sections: section A (14 items) inquired about the mother's religious practices; section B (10 items) inquired of the mother her husband's religious practices. However, something more than the concrete practice of religion was involved in responding to the items of part 1. Some of the items attempted to explore the individual's orientation toward religion, i.e., the intensity of faith. The following items should serve as examples of what was implied: "My faith has been a source of personal strength to me in everyday living"; "I have received real help with my problems from the other members of my religious faith"; "My faith has been strengthened as a result of having a retarded child."

Part 2 of the questionnaire was composed of 50 items reflecting attitudes of mothers about their retarded children. A number of these items were borrowed from the Parental Attitude Research Instrument (18), but revised

² A limited number of copies of the questionnaire are available and may be had upon request of the authors.

according to what were believed to be requirements of the sample. The PARI was considered too long and to require too high a level of verbal sophistication. In addition, the relevancy of many of the items of the PARI for the particular problem under study was questioned. Items were worded specifically to orient mothers to revealing attitudes about their retarded children, not about their other children or children in general. They could respond in one of four ways: strongly agree, mildly agree, mildly disagree, strongly disagree. A selected sample of the 50 items is as follows:

1. I frequently ask: Why did my child have to be slow?
2. I think an operation would cure or reduce my child's problem.
3. Frequently I give my slow child more love than he needs.
4. It is not my fault that our child is retarded.
5. My greatest mission in life is to care for my slow child.
6. I have not hesitated in telling our relatives what is wrong with our child.
7. Often my husband and I disagree on how to handle our child.
8. My husband and I are making plans for our child's future care and training.
9. I frequently become very impatient with my slow child.
10. In some ways my child is superior to normal children.
11. I do not discipline my child enough.
12. One of the most difficult things about taking care of a slow child is a woman feels that she can't get out.
13. I feel that my child was normal at birth but started having problems after he (or she) was older.
14. Having to care for a slow child has solved some of my more difficult problems.
15. I occasionally feel that something I have done has caused my child to be slow.
16. It sometimes bothers me to think of sending my child to an institution for slow children.
17. When I see all the normal children in the world, I frequently get disturbed that my child is not like the others.
18. God has chosen me to take care of a retarded child.
19. I tend to protect my child too much.
20. Even when I feel like showing my love for our slow child, I often have trouble.

This sample of items should make clear the effort to probe for and elicit feelings of guilt, anxiety and rejection on the part of mothers.

Subjects

The questionnaire was mailed to 125 mothers who had had contact with the Mental Retardation Clinic at St. Christopher's, a nondenominational hospital for children, from about September, 1957, to December, 1958. A letter describing the purpose of the questionnaire and the need for early return accompanied it.³ As a result of numerous follow-ups by mail and

³ The covering letter was sent out under the signature of one of the authors, Rev. Ralph L. Miller, a minister of the Presbyterian denomination.

CHILD DEVELOPMENT

telephone, 75 questionnaires were returned, 60 per cent of the maximal expected. Several reasons for failure to return the questionnaire were later discovered: lack of ability to follow directions (a few mothers were foreign-born and could not read English sufficiently well), change of address, and refusal to answer items for personal reasons. For these and other reasons, returns of the order of 60 per cent are typical in cases of such self-selected samples as the present. Considerably more elaborate studies in recent years have hardly exceeded this order of return; for example, *see* (20).

The group returning the questionnaire did meet the methodological requirement of being composed of about equal numbers of Catholics and non-Catholics. In addition, it was felt to be fairly homogeneous for education and socioeconomic status.

Table 1 gives some descriptive information about the mothers (*N* of 72, because the questionnaires of three mothers failed to supply needed identification material). There were 37 Roman Catholic mothers and 35 non-Catholic mothers. Of the non-Catholic mothers, 27 were Protestant and eight Jewish. Of the 27 Protestant mothers, 24 were white and three Negro. Information is given in Table 1 for numbers of children in the sample and sex of the child. Average formal education of the mothers was 11 years. The majority of fathers were employed in skilled and semiskilled trade.

TABLE 1
DESCRIPTIVE INFORMATION ON MOTHERS AND CHILDREN

Variable	CATHOLIC (<i>N</i> = 37)		PROTESTANT (<i>N</i> = 27)		JEWISH (<i>N</i> = 8)	
	Mean	Range	Mean	Range	Mean	Range
Age of mothers	33.5	24-46	32.0	24-51	37.2	29-46
Age of children	4.7	1-12	5.0	1-12	6.0	2-13
Number of siblings	2.5	0- 7	2.0	0- 6	1.6	1- 3
Most recent IQ of children	44.9	12-75	48.3	22-85	34.8	23-50

Scoring the Questionnaire

On part 1 of the questionnaire, responses were weighted as reflecting greater or lesser formalized religious observance. For example, the response "weekly" to the item, "I attend worship services . . ." received a weight of 5, the response "monthly" a weight of 3, the response "yearly" a weight of 1, and the response "not at all" a weight of 0. Several of the items were combined to yield a composite weighted score. For these, simple scoring matrices were developed.

On part 2 of the questionnaire, an attitude scale, the individual (R.L.M., a Protestant minister) who constructed the scale also prejudged responses as reflecting greater or lesser acceptance of the child by the mother. An effort to determine to what extent staff members agreed with the prejudged

responses was carried out by asking six members of the Clinic staff to fill out the questionnaire as they imagined an ideally acceptant mother might. Agreement of three or more of the staff with the prejudged rating occurred in 47 of the 50 items. On three remaining items, two of the staff agreed with the prejudged rating. None of these three items, however, on which agreement between staff and judge fell below 50 per cent was found to evoke differential response among the religious groups. While 50 per cent agreement between staff and judge can hardly be called "high," it seemed the best procedure to be as arbitrary as this in view of the state of knowledge of what constitutes maternal acceptance as reflected in attitude expression. Higher agreement concerning acceptance might have been found, but it seems more likely as a result of arbitrary compromise than obvious decision.

The 50 items of part 2 were broken down into five subscales, according as their content seemed to reflect the following factors: Overprotection (8 items), Discipline (10 items), Cause of Retardation (8 items), Acceptance of Diagnosis (9 items), and Mother's Self-Fulfillment (15 items).

Thus, from parts 1 and 2 of the questionnaire a total of eight scores were obtained: one of mother's religious practices, one of father's religious practices, five from subscales of part 2, and one from the total scale of part 2. In addition, each item of the full questionnaire was statistically evaluated to try to detect religious differences in response.

Clinical Measures

Clinical judgments of acceptance of the child by mother and adjustment of the child were obtained from the staff social worker and pediatrician.⁴ Acceptance ratings were based on absence of anxiety in the mother, feelings of adequacy with the child, and absence of hostility toward child or professional guidance. The impressions of the social worker were used primarily because: (a) the focus of her interviews with parents was more directly on their reaction to the child, and (b) she was a member of neither of the predominant religious groups compared in the study. Reported in a previous study (27) are levels of agreement between this worker's judgments and those of two psychologists who independently rated samples of social history protocols on which all judgments were based. In view of what was considered to be sufficiently high agreement among these individuals, it was deemed appropriate to use the social worker's judgments as a clinical measure of acceptance.

For 35 of the 72 mothers, the social worker's judgments were already available from the previous study. These evaluations had been made at the time of initial contact with the mother. For the remaining 37 mothers, however, the social worker was asked to make her ratings as of the time of the present study. The set of 72 ratings was correlated with the other variables

⁴ Special thanks are due Miss Helen Beck, social worker, and Dr. June Dobbs, pediatrician, for their evaluations of the mothers.

CHILD DEVELOPMENT

of the study (see Table 5). The set of 35 ratings made at initial contact was also correlated with the other variables of the study, but failed to show results differing significantly from those using the larger set of 72.

Table 2 lists, by religious background, the social worker's ratings of 35 mothers made at the time of initial contact and her 37 ratings made later. These ratings were then combined (N of 72) and compared with acceptance ratings made for the previous Zuk study. In the previous study, a significant correlation (chi square equalled 18.28, corrected for continuity; $p < .001$) was established between religious background and acceptance. In the present study, while a trend in the same direction was apparent, there was no significant correlation (chi square equalled 1.44, corrected for continuity; $p > .20$). The social worker rated as acceptant 19, or 51 per cent, of 37 Catholic mothers; 10, or 37 per cent, of 27 Protestant mothers; and two, or 25 per cent, of eight Jewish mothers.

Inspection of the 35 ratings carried over to the present study suggests the possibility of some bias in the return of the questionnaires. There was a comparatively higher percentage of return from mothers judged by the social worker to be nonacceptant. In the previous study, the ratio (based on an N of 39) of acceptant to nonacceptant Catholics was approximately 2:1. But in the 35 ratings carried over to the present study (i.e., mothers who returned the questionnaire), the ratio (based on an N of 18) was 1:1. In the previous study, the ratio (based on an N of 28) of acceptant to nonacceptant Protestants was about 1:5. But in the 35 ratings carried over, the ratio (based on an N of 13) was about 1:2.

Significant religious differences were not apparent in the pediatrician's ratings of the child's adjustment, although it is of some interest that there was a slight tendency to rate Catholic children toward the lower end of the scale. Of the Catholic children, one was rated excellent in adjustment, three

TABLE 2
SOCIAL WORKER'S EVALUATIONS OF MATERNAL ACCEPTANCE

	E V A L U A T I O N S							
	"Old" ($N=35$)		"New" ($N=37$)		Present Study ($N=72$)		Zuk Study ($N=76$)	
	A*	NA*	A	NA	A	NA	A	NA
Catholic Mothers	9	9	10	9	19	18	25	14
Protestant Mothers	4	9	6	8	10	17	5	23
Jewish Mothers	0	4	2	2	2	6	0	9

NOTE.—Included for comparison are social worker's "old" (carried to present from a previous study) and "new" (obtained for present study) evaluations of mothers' acceptance of child, and, also, evaluations made on mothers ($N=72$) who returned the questionnaire (present study, "old" plus "new") and on mothers ($N=76$) who composed sample of prior Zuk study (27).

* A = Acceptant; NA = Nonacceptant.

good, 21 fair, and 12 poor. Of the non-Catholic children, two were rated excellent in adjustment, five good, 21 fair, and seven poor.

The lack of significant correlation ($r = .21$; see Table 5) between the pediatrician's and social worker's judgments suggests that nonoverlapping criteria were being utilized for the judgments.

The Factor Analysis

A factor analysis (see Table 6) of the variables employed in the study was carried out under direction of one of the authors (F. K.). A modification of the group method approximating principal components was used (22). Special care was taken to limit the weighting of the dichotomous religious variables so as not to allow them to overdetermine the structure, since our primary interest was in the structure emerging from the acceptance and religious involvement variables.

The table of intercorrelations (see Table 5) revealed certain clusters of variables which tended to be positively correlated. Table 5 was reduced to its major components by the method of factor analysis described above. The factors seemed sufficient to account satisfactorily for the observed intercorrelations. The factor loadings (orthogonally rotated) of each variable are given in Table 6.

In the present case, it was inferred that factor I was characterized by acceptance of the retarded child and factor II by religious involvement. These factors were found to be more independent than anticipated. However, it was found that certain aspects of acceptance (namely, the variables, Overprotection and Discipline) tended to be related to religious involvement.

RESULTS

Split-half reliability coefficients of sections A and B of part 1 were .91 and .92, respectively. Table 3 shows mean scores obtained from sections A and B. Catholic mothers reported themselves and their husbands as more intensive in formalized religious observance than Protestant or Jewish mothers. Significant religious differences⁵ were found in this direction on three of the 24 items on part 1, with differences on a fourth item approaching the 5 per cent level. Catholic mothers, to a greater degree than Protestant mothers, reported that they were:

1. More faithful in church attendance;
2. More loyal to their parents' religious training;
3. More prone to attend church when everything was going well, rather than just in time of crisis;
4. More consistent in their practice of prayer.

Split-half reliability coefficients of the subscales of part 2 were .58 (Overprotection), .54 (Discipline), .72 (Cause of Retardation), .67 (Acceptance

⁵ The p levels were determined by the formula given in Walker and Lev (23, pp. 77-78).

CHILD DEVELOPMENT

of Diagnosis), and .64 (Mother's Self-Fulfillment). Table 3 also shows mean scores obtained on these subscales. Clear differences among the religious groups are not apparent, although the sums of mean subscale scores do reflect differences in the predicted direction. For the Catholic group, the sum is 26.2; for the Protestant group, 23.9; and for the Jewish group, 24.2.

TABLE 3
AVERAGE SCORES OF MOTHERS, BY RELIGIOUS GROUP, ON QUESTIONNAIRE

	<i>Catholic</i> (<i>N</i> = 37)	<i>Protestant</i> (<i>N</i> = 27)	<i>Jewish</i> (<i>N</i> = 8)	<i>Maximum</i> <i>Score Possible</i>
<i>Part 1 (Sections A & B):</i>				
Mother's Religious Observance ..	45.0	39.0	20.0	58
Father's Religious Observance ..	24.0	15.0	11.0	34
Sums	69.0	54.0	31.0	92
<i>Part 2 (Subscales):</i>				
Mother's Self-Fulfillment	7.2	7.4	7.0	15
Discipline	6.0	5.4	5.8	10
Overprotection	3.0	2.4	1.6	8
Acceptance of Diagnosis	5.7	4.8	6.5	9 ^a
Acceptance of Cause	4.3	3.9	3.3	8
Sums	26.2	23.9	24.2	50

Table 4 reveals some measure of association between parts 1 and 2 of the questionnaire. Taken by religious groups, individuals scoring "high" and "low" on part 1 are compared as to their scores on the subscales of part 2. "Highs" and "lows" of the Catholic and Protestant group were split at the upper and lower 27 per cent level. The Jewish group, because of its small size, was split 50-50. Significant differences within individual subscales are not apparent, but on the subscales taken together significant differences in the predicted direction occur: "high" Catholics scored higher on all subscales than "low" Catholics ($p < .03$), and "high" Catholics scored higher on all subscales than "high" Protestants ($p < .03$).

Significant religious differences were found on only three items of part 2. On each of these, however, Catholic mothers more frequently gave the pre-judged response reflecting greater acceptance of the child. More frequently than Catholic mothers, Protestant mothers stated that:

1. Their greatest mission in life was to care for their retarded child;
2. They more often became impatient with their retarded child;
3. Medical care could bring their child back to normal.

Taking all 50 items of part 2 together, the average percentage of "acceptant" response of the Catholic group was 26.5 per item (with a range of 2 to 91 per cent). The average percentage of "nonacceptant" response by the Catholic group was 21.6 per item (with a range of 0 to 50 per cent). The average percentage of "acceptant" response by the Protestant group was

TABLE 4

COMPARISONS OF SCORES ON PART 2 OF QUESTIONNAIRE OF "HIGH" AND "LOW" SCORERS ON PART 1, BY RELIGIOUS GROUP

Variables	CATHOLIC		PROTESTANT		JEWISH	
	"Highs" (N=10)	"Lows" (N=10)	"Highs" (N=8)	"Lows" (N=8)	"Highs" (N=4)	"Lows" (N=4)
<i>Part 2 (Subscales):</i>						
Mother's Self-Fulfillment	7.9	7.1	7.6	7.3	7.7	4.8
Discipline	6.7	5.9	6.2	5.3	6.7	4.8
Overprotection	4.1	2.6	3.1	2.0	1.1	2.0
Acceptance of Diagnosis	6.3	5.4	5.0	5.0	5.2	7.8
Acceptance of Cause	4.8	3.5	4.4	4.4	3.2	3.5
Sums	29.8	24.5	26.3	24.0	23.9	22.9

24.1 per item (with a range of 4 to 82 per cent). The average percentage of "nonacceptant" response by the Protestant group was 23.5 per item (with a range of 0 to 63 per cent). The differences are slight but in the predicted direction. It is of interest also that the average percentage of omitted responses per item was somewhat greater for the Protestant group (4.8 per cent with a range of 0 to 19 per cent) as compared with the Catholic group (4.1 per cent with a range of 0 to 27 per cent).

Some of the items composing part 2 approximate the wording of items that have been used in other studies of parents of handicapped children. For example, Ray (11) reported that 70 per cent of his sample of mothers of cerebral palsied children replied in the affirmative to the question: "Do you think God has given you a handicapped child as a 'cross to bear'?" In the present study, 61 per cent of the Catholic mothers (based on an *N* of 37), 29 per cent of the Protestant mothers (based on an *N* of 27), and 25 per cent of the Jewish mothers (based on an *N* of 8) replied in the affirmative to the item: "Having a slow child is the cross I must bear." Nineteen per cent of the Protestant mothers made no response to this item, but only 3 per cent of the Catholic mothers did not respond.

Boles (2) found that 68 per cent of his sample of mothers of cerebral palsied children and 70 per cent of his mothers of nonhandicapped children answered in the affirmative to the item: "I believe that if one puts in enough effort any condition can be changed." To an item in the present study, "I think my child will outgrow his problem," 45 per cent of the Catholic mothers, 52 per cent of the Protestant mothers, and 24 per cent of the Jewish mothers replied in the affirmative.

The importance of phrasing of items is highlighted by contrasting the responses to the item, "Having a slow child is the cross I must bear," (see above) with those to the item, "My greatest mission in life is to care for my slow child." On the latter, 89 per cent of the Protestant mothers, 61

CHILD DEVELOPMENT

per cent of the Catholic mothers, and 62 per cent of the Jewish mothers agreed. Although both items on the surface appear to be tapping the same attitude, a religious difference seems to have sprung up. One may conjecture that the Catholic group responded more positively to the phrase "cross to bear," the Protestant group to the phrase "mission in life."

Table 5 presents the results of a Pearson correlational analysis of the variables employed in the study. The correlation of section A with section B of part 1 was .48. The intercorrelations among the subscales of part 2 ranged from .23 (Overprotection with Acceptance of the Diagnosis) to .50 (Acceptance of Cause with Acceptance of Diagnosis). The correlations of scores reflecting intensity of mothers' religious observance with religious background varied from $-.60$ (Jewish) to $-.05$ (Protestant) to .42 Catholic). The range of correlations between intensity of fathers' religious observance and religious background was more restricted, $-.21$ to .27, but the directions were the same as those for mothers. Clinical judgments of the mother or child failed to correlate significantly with scores from parts 1 or 2. Correlations between the social worker's ratings of mothers' acceptance and parts 1 and 2 ranged from .05 to .22. Correlations between the pediatrician's ratings of adjustment of the child and parts 1 and 2 ranged from $-.13$ to .20.

As might be expected, the correlation between mothers' and children's ages was significant but slight, .28. A significant correlation (.34) was found between IQ and the pediatrician's ratings, which reflected a reasonable kind of bias in these ratings. It is rather more difficult to make sense of some of the other correlations that reached or exceeded the .05 level of significance: for example, the correlations between intensity of mother's religious practices and IQ of the child (.27), between the Acceptance of Diagnosis subscale and sex of child ($-.27$), or between the Discipline subscale and age of child ($-.24$), are hard to interpret. They must remain unexplained for the present or perhaps to be explained as chance occurrences.

Table 6 reports the results of a factor analysis of the variables of the study. Factor I has its highest loading on the subscales of part 2, whereas factor II has its highest loadings on sections A and B of part 1 and also on religious background. Very roughly, factor I may be defined as "Verbalized Acceptance of the Child," factor II as "Formalized Religious Involvement."

Perhaps the most interesting aspect of the factor analysis is the finding that two of the subscales of part 2 of the questionnaire, Overprotection and Discipline, have appreciable loadings on both factor I and II. To a lesser degree, the same holds true for the Acceptance of Cause subscale. Overprotection has a loading of .43 on factor I, .32 on factor II. Discipline has a loading of .52 on factor I, .40 on factor II. Acceptance of Cause has a loading of .67 on factor I, .25 on factor II.

These results reflect further on the relation between parts 1 and 2 of the questionnaire, between, that is, intensity of formalized religious observance

TABLE 5
INTERCORRELATIONS OF THE VARIABLES EMPLOYED IN THE STUDY

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Mother's Self-Fulfillment42*	.37*	.28*	.41*	.02	.04	.02	.04	-.04	-.14	-.19	-.08	-.01	.11	-.13	.11
2. Discipline26*	.33*	.44*	.21	.42*	.11	-.12	.01	-.11	-.24*	-.08	-.07	.12	.13	.06
3. Overprotection23*	.39*	.24*	.22	.21	-.08	-.20	-.12	-.02	-.10	-.08	-.03	.08	.05
4. Acceptance of Diagnosis50*	-.15	.02	.09	-.20	.16	-.17	-.04	-.01	-.27*	-.20	-.11	.09
5. Acceptance of Cause12	.20	.13	-.06	-.12	-.04	-.10	.12	-.07	.12	.13	.22
6. Mother's Religious Observance							-.48*	.42*	-.05	-.60*	-.21	-.26*	.16	.17	.27*	.11	.15
7. Father's Religious Observance27*	-.14	-.21	-.06	-.13	-.05	-.09	.10	.20	.17
8. Catholic02	-.09	.20	.08	-.02	-.11	.15
9. Protestant											-.16	.01	-.11	.01	.14	.21	-.07
10. Jewish21	.14	-.13	-.15	-.19	-.15	-.12
11. Mother's Age28*	.01	-.04	-.03	-.09	-.03
12. Child's Age09	-.02	-.16	-.04	-.20
13. Number of Siblings05	.13
14. Sex of Child13	.13
15. IQ of Child34*	-.10
16. Child's Adjustment (judged by pediatrician)27*
17. Mother's Acceptance (judged by social worker)21

* Correlations as great as or greater than $\pm .23$ are significantly different from zero at the .05 level. Correlations as great as or greater than $\pm .30$ are significantly different from zero at the .01 level.

CHILD DEVELOPMENT

and attitudes of parents of retarded children. Mothers who scored "high" in religious practice tended also to score "high" in verbally expressed attitudes reflecting acceptance, especially when factors of overprotection or discipline were involved.

TABLE 6
FACTOR ANALYSIS OF VARIABLES EMPLOYED IN STUDY

Variables	FACTOR LOADINGS	
	I	II
<i>Part 2 (Subscales):</i>		
Mother's Self-Fulfillment58	.13
Discipline52	.40
Overprotection43	.32
Acceptance of Diagnosis72	-.13
Acceptance of Cause67	.25
<i>Part 1 (Sections A and B):</i>		
Mother's Religious Observance	-.11	.75
Father's Religious Observance11	.62
<i>Other Variables:</i>		
Catholic07	.46
Protestant	-.16	-.13
Jewish15	-.56
Mother's Age	-.11	-.20
Child's Age	-.12	-.32
Number of Siblings01	.14
Sex of Child	-.25	.15
IQ of Child	-.11	.36
Child's Adjustment	-.10	.30
Mother's Acceptance (<i>N</i> = 35)15	.23
Mother's Acceptance (<i>N</i> = 37)07	.06

The following variables, in addition to those already described, had loadings greater than .25 on factor II: IQ of the child (.36), child's adjustment (.30), and child's age (-.31). There is no ready explanation for these findings. One additional variable had a loading of .25 on factor I: sex of the child (-.25). A slight tendency was found for mothers of female children to give responses on part 2 reflecting greater acceptance, consistent with Farber's (5) finding that the marital integration of parents was less affected by the presence of a retarded girl than of a retarded boy in the home.

DISCUSSION

A low but positive correlation was established between measures of religious background and maternal acceptance. The strength of relationship varied with the type of item on the attitude scale which was the measure of acceptance. Items on the attitude scale reflecting dispositions toward discipline of and overdependence in the child correlated more strongly

with religious background than other types of items. That is, mothers rating themselves more intense in their religious practices tended to express attitudes judged more acceptant, especially when these attitudes involved habits of discipline or overdependence.

Catholic mothers reported themselves and their husbands as more intense in religious practices than Protestant or Jewish mothers. On the attitude scale as a whole, Catholic mothers responded in a manner judged somewhat more acceptant of their retarded children. This finding gravitates in support of previous findings (5, 15, 27) that Catholic mothers are more acceptant than non-Catholic mothers and contrasts with Boles' (2) finding that Catholic mothers are more guilty, more anxious, and more socially withdrawn. The present finding suggests that maternal acceptance as reflected in verbalized attitudes, as well as in marital integration (5) and institutionalization (15), is influenced *in the same direction* by religious background.

The direction of influence was hypothesized as a result of differences in religious doctrine with respect to the problem of parental guilt. Catholic doctrine is explicit that parents should not feel guilt for bearing a retarded child; indeed, they are asked to accept the child as a special gift of God. Protestant and Jewish doctrine is neither so explicit nor absolving. Emphasis in the Jewish group particularly on intelligence as a special value constitutes an additional handicap in acceptance. Since it has often been noted in the clinical literature that guilt impedes acceptance of the handicapped child, it was believed that differences in acceptance might be detected as a result of the special Catholic injunction that parents not hold themselves personally at fault.

The clinical literature also states that there may be wide discrepancy between acceptance judged in different ways, as occurred for example in the failure to find correlation between measures of clinically-judged acceptance and verbalized acceptance on the attitude scale. The attitude scale, although found to be a fairly reliable measure when evaluated by split-half reliability coefficients of subscales, was not at all predictive when taken in relation to clinical judgment of acceptance as a validating criterion. It is known that attitudes and other behavior of parents toward their handicapped children may vary widely from one area to the next, and at times even be quite self-contradictory. Boles (2) found evidence for this in mothers of cerebral palsied children who proved unexpectedly realistic in appraising their children's mental ability, but "... when asked to estimate their children's future achievements, more basic unrealistic attitudes were revealed" (p. 211). Yet another kind of inconsistency was noted by Zuk (26) as a tendency of mothers to give information about their retarded children which portrayed the children in a better light developmentally than they were portrayed by their teachers. In so doing, the mothers were believed to be satisfying a frustrated need for their children to be more normal in intelligence than they actually were.

CHILD DEVELOPMENT

The recognition of handicap in their children provokes quite contradictory thoughts and feelings in parents, as Kanner has noted (7). He cited several cases of parents who could not forgive their children for failing to live up to expectations. These parents were intensely frustrated realizing that their children would not fulfill ambitions. How intense this frustration is may be envisioned in Bice's (1) report of two cases in which parents approached him to ask if they might discuss in group psychotherapy how they had once planned to kill their handicapped children. Clearly, certain thoughts and feelings of the parents can be outrightly hostile to the child. But the social setting in which parents live, the culture, frowns on this form of outright hostility, especially in the case of handicapped children. The hostility must, therefore, be kept under lock and key or be expressed in a more socially acceptable manner. It may even be directed against the child if suitably disguised, such as by applying pressure to perform beyond capacity. The psychoanalyst Rickman summarizes succinctly the psychological problem of parents of handicapped children when he asks consideration of the powerful conflict set up by "... man's seemingly eternal problem of love and hate for the same person—with ambivalence, which to many is an intolerable burden" (13, p. 125).

The intense conflict, frustration, and ambivalence of parents of retarded children are believed responsible for the difficulty in constructing *predictive* measures of acceptance of the child. This is not to say, however, that such measures do not provide a useful service in revealing clues as to the over-all condition of acceptance. The attempt to treat acceptance as an experimental variable clearly illustrates its many-faceted character. This variable, apparently so unitary and compact in clinical contexts, breaks up into parts and subparts when subjected to methodological treatment. Perhaps the greatest value of research on this variable is to demonstrate its multidimensional character, describe its various levels, and attempt to establish correlates among levels. Methodologically, treatment of acceptance in either molar or molecular interactions, as contrasted recently by Schaefer (17) in a re-evaluation of some maternal behavior concepts, would seem applicable to the general problem envisioned here. It is hoped that the present study has made some contribution in this direction.

SUMMARY

Maternal acceptance of retarded children was explored through means of a questionnaire which tapped attitudes toward the children and religious background. A low but positive correlation was found between measures of maternal acceptance and religious background. Mothers who rated themselves more intense in religious practices tended to verbalize attitudes judged more acceptant of their retarded children. Catholic mothers rated themselves and their husbands more intense in religious practices than non-Catholic mothers and also verbalized attitudes judged more acceptant.

Religious background correlated more positively with maternal acceptance when judgment of acceptance was based on attitude items involving dispositions toward discipline and overdependence.

The interpretation of the findings was believed to gravitate in similar direction to that of other findings which have recently been reported: namely, that Catholic mothers tend to be more acceptant of their retarded children than non-Catholics. Catholic doctrine is more explicit than non-Catholic in that parents consider themselves absolved of personal guilt in the event they bear a retarded or otherwise handicapped child. Since guilt has been assigned a role of major importance in impeding parental acceptance, it was believed that Catholics would be less blocked in acceptance than non-Catholics as a result of the dictum regarding guilt.

The presence of conflict, frustration, and ambivalence in mothers was believed to create special problems in validating a measure of acceptance from the attitude scale, although the measure was found to be reasonably reliable. These psychodynamics will probably continue to make it difficult to devise measures of acceptance based on expression of attitudes which will have much predictive capability. On the other hand, such measures can provide useful clues as to the condition of acceptance when used in comparison with other types of measures and observations.

REFERENCES

1. BICE, H. V. Group counseling with parents of the cerebral palsied. In *Psychological problems in cerebral palsy: a symposium*. Natl Soc. Crippled Child. & Adults, 1952.
2. BOLES, G. Personality factors in mothers of cerebral palsied children. *Genet. Psychol. Monogr.*, 1959, 59, 159-218.
3. BREITENBECK, G. *For parents of retarded children*. Liguori, Mo.: Liguorian Pamphlets, 1958.
4. EATON, J. W., & WEIL, R. J. *Culture and mental disorders: a comparative study of the Hutterites and other populations*. Free Press, 1955.
5. FARBER, B. Effects of a severely retarded child on family integration. *Monogr. Soc. Res. Child Developm.*, 1959, 24, No. 2.
6. HASTINGS, D. Some psychiatric problems of mental deficiency. *Amer. J. ment. Def.*, 1948, 52, 260-262.
7. KANNER, L. The emotional quandaries of exceptional children. In *Helping parents understand the exceptional child*. Langhorne, Pa.: Woods School, 1952. Pp. 21-28.
8. MILLER, R. L. Religious dimensions of maternal acceptance of the retarded child. Unpublished master's thesis, Princeton Univer., 1959.
9. MOORE, C. A study of the parental attitudes towards cerebral palsied children. Unpublished master's thesis, Univer. of Michigan, 1953.

CHILD DEVELOPMENT

10. PERLSTEIN, M. A. *The present-day status of the cerebral palsy problem.* Assn. Aid Crippled Child., New York City, 1947.
11. RAY, D. A study to develop a guide of education for parents of cerebral palsied children. Unpublished master's thesis, Univer. of Iowa, 1951.
12. RHEINGOLD, H. Interpreting mental deficiency to parents. *J. consult. Psychol.*, 1945, 9, 142-148.
13. RICKMAN, J. *Selected contributions to psychoanalysis.* Basic Books, 1957.
14. ROE, H. The psychological effects of having a cerebral palsied child in the family. Unpublished doctoral dissertation, Teach. Coll., Columbia Univer., 1950.
15. SAENGER, G. *The adjustment of severely retarded adults in the community.* Report to New York State Interdepartmental Health Resources Board, 1957.
16. SARASON, S. *Psychological problems in mental deficiency.* (2nd. Ed.) Harper, 1953.
17. SCHAEFER, E. S. A circumplex model for maternal behavior. *J. abnorm. soc. Psychol.*, 1959, 59, 226-235.
18. SCHAEFER, E. S., & BELL, R. Q. Development of a parental attitude research instrument. *Child Developm.*, 1958, 29, 339-361.
19. SCHEIMO, S. L. Problems in helping parents of mental defective and handicapped children. *Amer. J. ment. Def.*, 1951, 56, 42-47.
20. SEARS, R. R., MACCOBY, E. E., & LEVIN, H. *Patterns of child rearing.* Row, Peterson, 1957.
21. STONE, M. M. Parental attitudes to retardation. *Amer. J. ment. Def.*, 1948, 53, 363-372.
22. THURSTONE, L. L. *Multiple-factor analysis.* Univer. of Chicago Press, 1947.
23. WALKER, H., & LEV, J. *Statistical inference.* Holt, 1953.
24. WATERMAN, J. H. Psychogenic factors in parental acceptance of feeble-minded children. *Dis. nerv. System*, 1948, 9, 184-187.
25. WISHIK, S. M. *How to help your handicapped child.* Publ. Affairs Committee, 1955, No. 219.
26. ZUK, G. H. Autistic distortions in parents of retarded children. *J. consult. Psychol.*, 1959, 23, 171-176.
27. ZUK, G. H. The religious factor and role of guilt in parental acceptance of the retarded child. *Amer. J. ment. Def.*, 1959, 63, 139-147.

RECOGNITION OF PERCEPTUALLY AMBIGUOUS STIMULI IN GRADE SCHOOL CHILDREN

JURIS G. DRAGUNS*

Rochester State Hospital

and GIUSEPPE MULTARI¹

University of Rochester

Since Frenkel-Brunswik (4) introduced the concept of intolerance of ambiguity as a combined personality, perceptual, and cognitive variable, individual differences in readiness or reluctance to invest meaning into poorly structured stimuli have been employed as one of its operational indices (5, 9, 10, 13, 14). However, there are still no systematic data available on the manner in which these presumably constant and personality related response tendencies develop in childhood and adolescence. While a study by Korchin and Basowitz (6) has explored the effects of aging on perception and judgment of ambiguous visual stimuli, no such investigations have been conducted with children at various age levels. Several reports concerned with intolerance of ambiguity made use of preadolescent and adolescent subjects (9, 10, 14), but in none of these were there any age comparisons made.

More directly relevant to the problem at hand is the investigation by Dalla Volta (2) whose interest was focused on the ability of children between the ages of 4 and 12 to describe and reproduce pictorially various kinds of "naturally indistinct" stimuli, e.g., clouds. Dalla Volta's principal finding points to small age differences and to an unquestioning and presumably nonanxious attitude of his subjects in dealing with these expectedly and naturally vague parts of their environment. His report raises two questions appropriate to this investigation: Does this kind of behavior hold when a child is confronted with an artificially induced, unexpectedly ambiguous stimulus? Would age differences be more prominent if the subjects were

* Department of Psychology, Rochester State Hospital, 1600 South Avenue, Rochester 20, New York.

¹ Now at the University of Ottawa.

faced with "artificial indistinctness" induced by one of the many techniques of stimulus impoverishment?

Two other research reports throw some light on these issues. L'Abate (8) found, upon comparing the responses of children from kindergarten through 5th grade to a multiple choice picture story test and to some simple animal pictures, an increase in the number of "uncertainty expressions" and qualifiers as a function of age. Piaget and Stettler - von Albertini (12) reported that incomplete or overlapping geometrical figures tend to give rise among 4- to 6-year-olds to perception of "empirical configurations" such as houses, people, and utensils while 7- to 8-year-olds more veridically described such stimuli in terms of various geometrical and other abstract shapes. We interpret this finding as possibly indicating a greater tendency toward early—and maybe premature—meaningful structuring of visual stimuli at an early age. The results of both L'Abate and of the Swiss investigators would suggest that early age levels are characterized by quick and conflict-free recognition of ambiguous stimuli, unimpeded apparently by critical appraisal of the scant bases for such judgments. To explore further the relevance of age factors in children to the attribution of meaning to ambiguous stimuli is the object of the current study.

SUBJECTS

Thirty Ss each, consisting of an equal number of boys and girls, were selected randomly from various Catholic parochial schools in Rochester, New York,² at the following grade levels: 1st, 3rd, 5th, and 7th. These grades were chosen so as to reduce overlap in chronological age and still have a reasonably continuous picture of the behavior of children of various ages within the grade school range when confronted with our stimuli.

To explore the intellectual composition of our groups, an appropriate form of the Kuhlmann-Anderson intelligence test was group administered to a given grade level. The obtained means and standard deviations of IQs, summarized in Table 1 along with data on the Ss' age, show that the grade and sex groups are largely comparable in intelligence as no differences among them approach statistical significance.

MATERIALS AND PROCEDURE

The task selected to assess our Ss' meaningful structuring of ambiguous stimuli follows largely the "decision location" format employed in the study of recognition response level by Binder (1), Messick and Hills (10), Smock (13, 14), and others. In all of these, a series is presented starting with a

² Sincere appreciation is expressed to Rev. William M. Roche, Diocesan Superintendent of Schools, for permitting us to examine our subjects and to Sisters M. Beatrice, SSND, Gertrude Marie, SSJ, and Helen Therese, SSJ, for assistance in making classes available for our experiment.

TABLE I

MEANS AND STANDARD DEVIATIONS OF AGES AND INTELLIGENCE
QUOTIENTS IN GRADE AND SEX GROUPS

	AGE		INTELLIGENCE	
	M	SD	M	SD
<i>Grade I</i>				
Boys	6.9	0.61	105.2	7.5
Girls	6.7	0.36	109.4	4.1
Total	6.8	0.56	107.3	6.5
<i>Grade III</i>				
Boys	8.2	1.27	116.7	15.2
Girls	8.0	0.65	114.3	13.8
Total	8.1	0.78	115.5	14.8
<i>Grade V</i>				
Boys	9.8	0.64	103.1	8.3
Girls	9.8	0.65	99.1	9.1
Total	9.8	0.64	101.1	8.9
<i>Grade VII</i>				
Boys	12.6	0.74	100.7	7.7
Girls	12.6	0.42	106.8	9.1
Total	12.6	0.64	103.8	8.8

stimulus which, perceptually or cognitively, conveys minimal information to the subject. The series then proceeds, with appropriate graduations, to a presentation which permits unhampered recognition of the object or contains all of the information for the solution of the problem at hand. To construct such a series, various methods of obscuring, camouflaging, or incompletely presenting the stimulus have been employed. In our case, we followed a procedure applied by Voigt (15) in a different context. For our purposes, the following nine drawings were made: a sail boat, a bridge, a dog, a bowl of fruit, a house, a telephone, a group of trees, a violin, and a wheelbarrow, and then photographed with the objective progressively moved out of focus.³ In this manner, at least 12 pictures of each object were prepared. These were presented to 10 adult judges to see if their ranking of the pictures from unclear to clear would correspond to the degree of actual photographic blurring. On the whole this correspondence

³ The help of Miss Virginia Jeffrey Smith in providing the drawings for our experiment is gratefully acknowledged. We also thank Mr. Victor Atyas for his help in photographing our stimuli.

CHILD DEVELOPMENT

was high. The few photographs on which the judges' rankings were inconsistent were eliminated, but no other changes were made. Thus for experimental use, nine series of 12 pictures each were retained which ranged, in most cases, from the representation of a formless, blurred mass to a virtually unequivocally perceptible picture of a common object.

Each series of pictures was arranged in a booklet. These booklets were presented to the subjects in a random order to eliminate possible learning, habituation, or fatigue effects. The following instructions were given:

I have several booklets of pictures to show you. Some of the pictures in these booklets are kind of blurry. As we go through the booklets, I would like you to tell me when you feel you have an idea what the picture is. So keep in mind that it's okay to stop me at *any* point when you feel you have an idea what the picture is.

Each picture was shown to the subject for 10 seconds at the end of which time *E* turned the page, thus revealing the next picture. Presentation time was restricted to prevent perseverations and large variations in reaction time (which are not to be explored in this study), but at the same time to allow a fairly complete and unhurried examination of the card. With this procedure, all reactions occurred within the 10-second time limit, and wide variations in reaction time were not observed. The entire administration was completed within 10 to 20 minutes in most cases.

Presentation of each series yielded us one response with the following kinds of information: (a) the number of the picture, from 1 (at the vague end of the series) to 12, on which the *S* ventured to report seeing some sort of a concrete object; (b) whether or not the object seen was the one presented clearly on the final picture series. The former was used to compute the recognition level score; the latter was included in the total tally of correct or error responses, respectively. The *S*'s recognition level score

TABLE 2
SPLIT-HALF CORRELATION COEFFICIENTS FOR RECOGNITION
LEVEL AND ERROR SCORES

	Recognition	Errors
1st grade78	.27
3rd grade65	.68
5th grade39	.21
7th grade85	.37
Total73	

NOTE.—Pearson *r*s were computed for recognition level and Spearman *r*hos for errors.

was so derived for each series, and an individual's total score was found by averaging his nine obtained scores. The number of errors, only incidental to the main purpose of this investigation, was likewise computed for each individual.

To obtain an estimate of the reliability of Ss' recognition level and error scores, split-half correlation coefficients were computed by comparing their performance on the first five and last four series. As will be seen from the examination of Table 2, reliability is, on the whole, moderately high for the recognition responses of various grade groups as well as for the total population. Error scores, however, are much less consistent, it would seem.

RESULTS

Recognition level scores were investigated by means of a two-way analysis of variance for sexes (2) by grades (4). Table 3 summarizes the results of this analysis and reveals highly significant effects due to grade level ($p < .01$) while effects of both sex and sex-grade interaction are negligible. Comparison of mean grade level scores by means of t tests yields only one difference at the .05 level of confidence, that between the 1st and the 5th grade. Differences between 1st and 7th and between 3rd and 5th grades are of borderline significance ($p < .10$) in both cases.

Post hoc examination of recognition level means and standard deviations shown in Table 4 suggests a rather pronounced contrast between the two lower and the two higher grades, an impression which is confirmed by statistical analysis ($p < .05$).

Since distribution of error scores deviated markedly from normality, a distribution-free chi square analysis of variance, as described by Wilson (17), was applied to these data. This procedure, again, revealed significant grade effects at the .01 level, but no other reportable findings. Mann Whitney U tests performed to compare differences among various grade levels resulted in the following significant findings at the .05 level of confidence:

TABLE 3
ANALYSIS OF VARIANCE FOR RECOGNITION LEVEL IN SEX
AND GRADE GROUPS

	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Grade	3	96.24	18.43	.01
Sex	1	6.85	1.28	<i>ns</i>
Grade \times Sex	3	6.48	1.22	<i>ns</i>
Within Cells	112	5.27		

CHILD DEVELOPMENT

1st grade vs. 5th grade, 1st vs. 7th, and 3rd vs. 5th grade. Comparison of means and standard deviations of error scores as shown in Table 4, though possibly misleading in view of their skewed nature, points to a well-nigh continuous decline in the number of errors from our youngest to our oldest group while a tendency to a later object recognition increases with age. On

TABLE 4
MEANS AND STANDARD DEVIATIONS OF RECOGNITION LEVEL AND
ERROR SCORES IN GRADE AND SEX GROUPS

	RECOGNITION		ERRORS	
	M	SD	M	SD
<i>Grade I</i>				
Boys	3.0	1.1	8.4	1.1
Girls	4.8	2.6	7.2	3.1
Total	3.9	2.1	7.8	1.9
<i>Grade III</i>				
Boys	4.6	2.4	5.9	2.7
Girls	4.4	2.9	5.3	3.0
Total	4.5	2.7	5.6	2.9
<i>Grade V</i>				
Boys	7.6	0.9	4.6	1.7
Girls	8.0	2.9	4.5	2.1
Total	7.8	1.9	4.6	1.9
<i>Grade VII</i>				
Boys	6.7	2.4	4.7	2.6
Girls	7.2	1.9	5.1	1.9
Total	7.0	2.2	4.9	2.1

this basis and on logical grounds, one would expect a sizeable negative correlation between errors and recognition level; the earlier an attempt at object naming, the more of a chance there is for making a mistake in view of the incomplete information available. This expectation is, by and large, borne out by the data in Table 5. More noteworthy perhaps are the results of correlating these two mutually associated measures with intelligence. While recognition scores consistently show a negligible correlation with the IQ level, there is, in some grades at least, a significant negative correlation of errors with intelligence. This finding parallels that of Smock (14) who reported no relation between intelligence and the S's first response to an analogous task while a significant correlation was obtained between intelligence and the first correct response.

TABLE 5

CORRELATION COEFFICIENTS AMONG RECOGNITION LEVEL, ERRORS,
AND INTELLIGENCE QUOTIENTS

	Recognition Level with Errors	Recognition Level with Intelligence	Errors with Intelligence
1st grade	-.45**	.18	-.44*
3rd grade	-.74**	.12	-.42*
5th grade	-.51**	.06	-.32
7th grade	-.69**	.12	-.17
Total09	

NOTE.—Spearman *rhos* were computed, instead of Pearson *rs*, for all correlations with error scores.

* Significant beyond .05 level.

** Significant beyond .01 level.

DISCUSSION

Within a limited age span, our results seem to point to an increased cautiousness in structuring vaguely seen stimuli as a function of either higher age or greater educational attainment and to a concomitant decrease in the number of errors. The absence of sex differences may be worth noting in view of L'Abate's reporting such differences in his previously cited project (8). However, since the above author was investigating variations in verbal behavior while we were concerned with perceptual recognition, we can merely report this possible discrepancy without speculating at this time on the generality or specificity of the findings which gave rise to it.

Our positive results should be interpreted with caution. We do not yet know whether a continuous trend toward more conservative and more accurate judgments would hold true for the entire span from preschool age through adulthood or whether the small and statistically insignificant reversal to lower recognition level scores in the 7th grade may presage the articulation of such a tendency in adolescence. Furthermore, the results obtained could be attributed to either age or educational level or a combination of both. Perhaps, they may be regarded as a function of all relevant effects, educational and otherwise, which have impinged upon an individual up to the time of our experiment.

With the above limitations in mind, our data could be explained from two different, but overlapping, points of view. Within the developmental framework, our younger Ss' tendency toward early and often incorrect identification of picture contents might be interpreted, in Werner's (16) terms, as an instance of syncretism, i.e., a lack of segregation between internally and externally determined experiences. If we adopt the microgenetic

approach (3, 7, 15) and assume a continuous perceptual progression from the initial presentation to final structuring, the process of recognition may be viewed as a gradual balancing of subjective determinants and objective clues. The S's choice is then between extending the exploration of the picture—i.e., to use Piaget's terms, (11) "accommodating" himself to its emerging properties—or cutting this search short by "assimilating" the percept to the objects of his previous experience. The point on a continuum from indistinctness to clarity at which such decisions are made constitutes our basic measure and, on the basis of our findings, seems to differ with age and/or educational level.⁴ Our findings, however, in no way refute the possibility—and the relatively wide range of differences within our age groups indirectly tends to support it—that personality factors may contribute, along with developmental ones, to determining our S's scores. Our interpretation then need not dispense with the assumption either explicitly or implicitly made in many studies of intolerance of ambiguity, viz., that perceptual—and other—kinds of ambiguity are, to varying degrees, productive of anxiety (4, 14). Italian research (2, 18) and our own episodic impressions, however, suggest that children's reactions to poorly structured parts of the environment are free of at least overt discomfort.

In this, there seems to be a certain contrast between our data and those obtained by Korchin and Basowitz (6) on the judgment of ambiguous stimuli in the course of aging. While their experimental materials and procedures are not fully comparable to ours, they noted appreciable differences in responses to situational ambiguity—inherent in the incomplete structuring of the purpose and task of the experiment—in contrast to minimal differences in the readiness to judge, and the correctness of judging, actual ambiguous experimental materials. Integrating Dalla Volta's (2) and Zecca's (18) results with ours, we see a suggestion of a reverse tendency in children; while the ambiguities of the experimental task are accepted unreflectively and nonanxiously, stimulus recognition as such gives rise to perceptible age differences. Future research may investigate the hypothesis that development in childhood is characterized by changes in responding to stimulus ambiguity, but not to situational ambiguity.

Finally, it may be tempting if premature to try integrating our results with the personality oriented literature on intolerance of ambiguity. Thus,

⁴ A recent report by J. Drösler and W. F. Kuhn, which came to our attention after preparing this article, points out, on the basis of perceptual and psychophysical experiments, a greater tendency toward symmetry, structuring, and closure in children as compared with adults and consequently parallels to a certain extent our findings. See Drösler, J., & Kuhn, W. F., *Ein experimenteller Vergleich der visuellen Wahrnehmung von Kindern, Schizophrenen und Alkoholikern mit der tachistoskopischen Wahrnehmung normaler Erwachsener. XVI International congress of psychology. Individual papers. Summaries.* Bonn: German Society of Psychology, 1960, 1, 1-1, 2.) Less directly related but still relevant are the results of F. Winnefeld (*Gestaltauffassung und Umgestaltung in genetischer Sicht, Z. exp. angew. Psychol.*, 1959, 6, 589-602) who reports increases in the ability to find hidden figures and to reverse figure and ground as a function of increasing age. This appears to be in line with a more object determined response to ambiguous stimuli observed by us.

tendencies by anxious individuals toward an accelerated appearance of attempts at labeling an object and imbuing it with meaning (14) may conceivably be conceptualized as regression to modes of responding characteristic of earlier stages of development. Difficulties, however, arise in pursuing this line of reasoning as the tendency of avoiding commitment to material which is open to several interpretations has also been found among anxious Ss (5, 13, 14). Perhaps, the latter kind of response tendency may represent direct effects of anxiety while the former may be akin to a defense against it. Again, future investigations only can clarify this, at the present time, quite speculative point.

SUMMARY

Children of both sexes in four different grades were compared as to their readiness to attribute meaning to pictures differing in the amount of photographic blurring. Consistent if small differences among grades suggested a tendency toward a greater amount of cue accumulation before making such judgments as a function of age and/or grade level. The number of errors on this task likewise declined at the higher grade levels. Relevance of these findings within a developmental framework was discussed, and possible parallels as well as differences were noted upon the comparison of our data with the performance of aged individuals and of anxious subjects on similar tasks.

REFERENCES

1. BINDER, A. Personality variables and recognition response level. *J. abnorm. soc. Psychol.*, 1958, 57, 136-142.
2. DALLA VOLTA, A. Rappresentazione dell'indistinto a sviluppo del bambino. *Minerva Pediat.*, 1951, 3, 44-79.
3. FLAVELL, J. H., & DRAGUNS, J. A microgenetic approach to perception and thought. *Psychol. Bull.*, 1957, 54, 197-217.
4. FRENKEL-BRUNSWIK, E. Intolerance of ambiguity as an emotional and perceptual variable. *J. Pers.*, 1948, 18, 104-141.
5. HAMILTON, V. Perceptual and personality dynamics in reactions to ambiguity. *Brit. J. Psychol.*, 1957, 48, 200-215.
6. KORCHIN, S. J., & BASOWITZ, H. The judgment of ambiguous stimuli as an index of cognitive functioning in aging. *J. Pers.*, 1956, 25, 81-95.
7. KRAGH, U. *The actual-genetic model of perception-personality*. Lund: C. W. K. Gleerup, 1955.
8. L'ABATE, L. Sanford's uncertainty hypothesis in children. *Etc. Rev. gen. Semant.*, 1957, 14, 210-221.

CHILD DEVELOPMENT

9. LEVITT, E. E. Studies in intolerance of ambiguity: I. The decision location test with grade school children. *Child Developm.*, 1953, 24, 263-268.
10. MESSICK, S., & HILLS, J. R. Objective measurement of personality: cautiousness and intolerance of ambiguity. *Educ. psychol. Measmt.*, in press.
11. PIAGET, J. *The origins of intelligence in children*. International Universities Press, 1952.
12. PIAGET, J., & STETTLER - VON ALBERTINI, B. Recherches sur le développement de perception XIX. Observation sur la perception des bonnes formes chez l'enfant par actualization des lignes virtuelles. *Arch. Psychol* (Genève), 1954, 34, 203-242.
13. SMOCK, C. D. The influence of psychological stress and the "intolerance of ambiguity." *J. abnorm. soc. Psychol.*, 1955, 50, 177-182.
14. SMOCK, C. D. The relationship between "intolerance of ambiguity," generalization, and speed of perceptual closure. *Child Developm.*, 1957, 28, 27-36.
15. VOIGT, J. Die Aktualgenese in der psychologischen Diagnostik. *Psychol. Beitr.*, 1956, 2, 586-636.
16. WERNER, H. *Comparative psychology of mental development*. International Universities Press, 1957.
17. WILSON, K. V. A distribution-free test of analysis of variance hypotheses. *Psychol. Bull.*, 1956, 53, 96-101.
18. ZECCA, G. Il comportamento grafico del bambino nelle prime rappresentazioni degli indistinti naturali. *Arch. Psicol. Neurol. Psichiat.*, 1956, 12, 736-754.

QUANTITY CONCEPTIONS IN JUNIOR AND SENIOR HIGH SCHOOL STUDENTS¹

DAVID ELKIND*

UCLA School of Medicine

In Piaget's (11) experiments dealing with the development of quantity conceptions in children, he found that abstract responses to mass, weight, and volume appeared in a regular sequence that was related to age. Abstract responses to mass appeared (i.e., in 75 per cent of the subjects tested) by ages 7 to 8, abstract responses to weight appeared at ages 9 to 10, and abstract responses to volume by ages 11 to 12. A systematic replication of one of Piaget's experiments (6) confirmed Piaget's ages for mass and weight but failed to confirm his age for volume which was conceived abstractly by only 27 per cent of 11- to 12-year-old American children. The present study seeks to extend the replication to 12- to 18-year groups and to determine the influence of age, sex, and IQ on the attainment of abstract conceptions of quantity in adolescents.

According to Piaget, the criterion of an abstract conception of quantity is the child's judgment of its conservation, its sameness despite perceptual change. For example, in one of Piaget's experiments he showed his subjects

* Department of Psychiatry, School of Medicine, University of California Medical Center, Los Angeles 24.

¹ This investigation was supported (in part) by a PHS Small Grant M-3466 from the Department of Health, Education, and Welfare, U. S. Public Health Service; the author was then at Wheaton College.

CHILD DEVELOPMENT

two identical clay balls, one of which was subsequently rolled into a "sausage." Those children who judged the mass, weight, or volume of the "sausage" as conserved (the same) demonstrated an abstract conception of quantity. The same technique and criterion were used in the present study.

METHOD

Subjects

Four hundred and sixty-nine junior and senior high school students attending the Norton High School in Norton, Massachusetts, were tested.² The majority of children came from lower-middle class families. The mean age and IQ for boys and girls at each grade level are given in Table 1.

Tests and Procedure

The tests for conservation were given as group³ tests, and the students wrote their answers in blue books. Materials for all tests were the same: two identical clay balls (1½ inches in diameter) and a small scale. The tests were always given in the following order.

1. *Test for the conservation of mass.* *E* began by explaining that the two balls were identical in every way, that there were no tricks in the experiment, and that he was simply repeating an experiment originally performed with elementary school pupils. Several students were asked to verify by means of the scale that the two balls weighed the same. *E* asked any student who still had doubts to voice them. Questions raised permitted *E* to explain further that the balls were made of the same kind of clay and that "same amount" meant to the degree of accuracy permitted by the scale and did not mean absolute equivalence.

Once the students agreed that the balls were the same, *E* asked: (a) "Do the balls both contain the same amount of clay?" (identity question); (b) "Suppose I made one of the balls into a sausage, would the two pieces of clay still contain the same amount of clay?" (prediction question). *E* then actually rolled one of the balls into a sausage and asked: (c) "Do they both contain the same amount of clay now?" (judgment question); (d) "Explain your answers" (explanation question).

2. *Test for the conservation of weight.* *E* rolled the "sausage" back into a ball and then repeated the whole procedure described above save for the use of "weight" instead of "amount."

3. *Test for the conservation of volume.* *E* again rolled the "sausage" back into a ball and repeated the procedures for mass and weight, except that the terms "volume" and "same room or space" were substituted for the previously used quantity concepts.

² The writer is indebted to Mr. Henri A. Yelle, Principal of the Norton High School, and to the teachers and students whose friendly cooperation made the study possible.

³ In Piaget's (12) experiments and in the first replication (6) with elementary school children, the tests were administered individually.

Controls

Objections were raised by Baldwin (2) against Dennis' (3) finding of animism in college students. The brunt of these objections was that verbal misunderstandings and verbal ambiguities made interpretation hazardous. In the present study a number of controls were introduced to minimize the verbal misunderstanding and ambiguity: (a) the preliminary procedures of weighing the balls and of answering questions regarding the experiment tended to reduce misunderstandings; (b) the four types of question asked for each concept—identity, prediction, judgment, and explanation—gave some measure of the reliability of responses; (c) finally, as a special check on the conservation of volume test (which proved most difficult for children in the original and replication study), students were asked what would happen if the ball and the sausage were put into identical glasses filled equally high with water.

Scoring

A test was considered passed only if the subject correctly identified, predicted, judged, and explained the conservation of a quantity.

RESULTS

Controls

The control tests indicated that students who failed the conservation tests did so because of inadequate conceptions and not because of verbal misunderstanding.

All of the students who failed the volume test agreed (answered "Yes" to the identity question) that initially the balls were the same in volume. But they predicted, judged, and explained that the volume of the clay changed with the change in its shape. This was true for the same students who on the previous pages had written that mass and weight were conserved because either: (a) nothing was added or taken away; or (b) changing the shape did not change the amount; or (c) what the sausage lost in width it gained in length and therefore was the same as the ball.⁴ Thus, the students clearly understood the conservation questions with regard to mass and weight, but failed to generalize their judgments, predictions, and explanations to the conservation of volume which they treated as an entirely different problem. Kay (16-5) wrote, for example, that mass and weight were conserved because nothing was added or taken away but to the volume question she wrote, "The molecules may be more compressed in one object (the ball) than in the other (the sausage); although it (the sausage) has the same number of molecules and the same weight, its volume is not the same."

The question about the water levels, introduced as a control, revealed a fact significant in itself. Over two thirds (70 per cent) of the students

⁴ These are the same three explanations given by the elementary school children.

CHILD DEVELOPMENT

said that the water levels would rise because of the weight (often described as impact or pressure) of the submerged objects. Gla (15-4) "The water with the ball will rise more because the impact (*sic*) of the ball is greater because it is more condensed and it will rise more."

Such examples could be multiplied, but even these are sufficient to show that students who failed to judge volume conservation did so, not because of verbal misunderstanding, but rather because they failed, in Piaget's terms, to dissociate their subjective sensorimotor conceptions (impact, compression) of weight and volume from their objective (molecules), logico-mathematical conceptions.

Percentage of Subjects Judging Conservation

Of the 469 junior and senior high school students tested, 87 per cent had abstract conceptions of mass and weight, but only 47 per cent had an abstract conception of volume. A test of the difference between the three percentages gave an χ^2 of 254.2 which is significant beyond the .001 level. Table 1 shows that, according to the usual 75 per cent criterion for placing a test at a given year level, the volume test is too difficult for all but the

TABLE 1
PERCENTAGE AND MEAN IQ OF STUDENTS, CLASSED BY SEX AND AGE,
HAVING ABSTRACT CONCEPTIONS OF QUANTITY

Sex	N	Mean Age	Mean IQ	QUANTITY CONCEPTION		
				Mass	Weight	Volume
M	56	12.6	94.6	79	71	38
F	66	12.6	99.2	79	86	26
M	46	13.6	89.6	78	74	43
F	48	13.6	100.9	90	92	29
M	28	14.6	99.6	93	96	68
F	45	14.6	107.9	93	93	40
M	31	15.5	95.3	97	90	58
F	41	15.7	101.9	98	93	39
M	29	16.4	100.0	83	90	72
F	43	16.6	107.9	86	86	58
M	14	17.7	109.9	100	100	79
F	12	17.7	108.3	100	95	68
T O T A L S						
M	204	15.1	96.5	85	83	54
F	265	15.1	104.3	89	90	40
M & F	469	15.1	100.4	87	87	47

DAVID ELKIND

oldest age group. This result, like that of the first replication, fails to confirm Piaget's age (years 11 to 12) for the abstract conception of volume.

Variables Influencing Volume Conservation

Age. As shown in Table 1, there was a regular increase with age in the percentage of students who had abstract volume conception.

Sex. Table 1 shows that for each age level the percentage of boys was consistently higher than the percentage of girls who had an abstract conception of volume. For the two groups as a whole the difference gave an χ^2 of 7.9 which is significant at the .01 level.

IQ. The point biserial coefficient for the correlation of IQ (Kuhlmann-Anderson) with passing the volume test was .31, significant beyond the .01 level. This finding agrees with the results of other studies (4, 5, 6) which have also shown a low but positive correlation of IQ with measures of conceptual development.

DISCUSSION

The results of the present study raise two interrelated questions: (a) Why do significantly more students attain abstract conceptions of mass and weight than attain the abstract conception of volume? (b) In what respect are age, sex, and IQ related to volume conservation? Both questions can be answered within the framework of Piaget's genetic psychology.

In the course of more than 40 years of continuous research on mental development Piaget has evolved a genetic theory of intelligence, adaptive thinking and action. He derives, in a necessary sequence of age-related stages, the abstract intelligence of adults from the sensorimotor coordinations of infants. It is a nature theory in the sense that the stages—sensorimotor (0 to 2 years); preoperational (2 to 7 years); concrete operational (7 to 11 years); formal operational (11 to 15 years)—are assumed to reflect maturational changes in the forms of thinking. It is also a nature theory in the sense that the gradual attainment of particular conceptions is assumed to reflect the influence of physical and social environment on the contents of thought.

Unlike reinforcement learning theorists, Piaget holds that there is no real separation between drives and cognitive structures: "À tous les niveaux, les intérêts et les besoins qui son intervenus dans ces apprentissages étaient eux mêmes solidaires des structures déjà construes ou en voie de construction" (7, p. 30). For Piaget the appearance of new structures gives rise to new drives to exercise those structures. Such exercise is self-reinforcing through the production of function pleasure. In the course of being exercised, novel stimulation leads to the differentiation of new structures which in turn give rise to new drives in a continuous progression.

The attempt by the organism to use the stimulation of the environment to develop and conserve its structures, Piaget speaks of as assimilation. The

modification of structures due to environmental influence, he speaks of as accommodation. Any particular cognitive system attains its final stage of development when assimilation and accommodation are in equilibrium. The two processes are in equilibrium when novel stimulation can be used by the system without either disrupting its activity or necessitating a structural differentiation. According to Piaget, the attainment of the equilibrium stage is the underlying dynamic of all intellectual growth.

Applied to quantity conceptions, Piaget's theory holds that the child is ready⁵ for attaining the abstract volume conception at ages of 11 to 12. The child is ready in the maturational sense because the conceptualization of volume requires only concrete operations (internalized actions) which are present in most children by the age of seven (5, 6, 11, 12). The child is also ready in the experiential sense that he has had sufficient object contacts to form abstract conceptions of mass and weight which are the structural prerequisites for the attainment of the volume conception (6, 12). On the basis of his findings Piaget (12) assumed that this operational and structural readiness was sufficient for the attainment of the volume conception by a majority of 11- to 12-year-old children. His theory, on the contrary, suggests that a majority of young people should not, as in the present study they did not, attain an abstract conception of volume.

Piaget's theory proposes two kinds of events which when taken in combination can account for the results of this investigation. According to Piaget, the age at which a young person is ready to attain the volume conception is also the age at which he is developing formal mental operations. Whereas concrete operations (7 to 11 years) are concerned with classifying, relating, and quantifying the immediate environment, formal operations are concerned with constructing systems and theories. Put simply, concrete operations are concerned with immediate reality, formal operations with possibility (13).

Because of the unity of operations and drives, the appearance of formal operations means the appearance of new interests and horizons:⁶ "The adolescent is no longer content to live in the interindividual relations offered by his immediate surroundings or to use his intelligence to solve problems of the moment" (13, p. 341). The result of the new interests and motivations arising from formal operations is necessarily an attenuation in the drive for conceptualizing the immediate physical environment. This attenuation of drive for inductive conceptualizing therefore decreases the possibility of children spontaneously discovering the conservation of volume.

Intimately bound up with the appearance of formal operations is a second event which affects the early adolescent's conceptualizing activity. This second event is the adoption, beginning around ages 11 to 12, of

⁵ This is not Piaget's term but economically describes the child's state.

⁶ This may be one of the reasons for the well established decline of interest in school at about the same age (9, 10, 14). Perhaps, the school curriculum does not take sufficient account of the early adolescent's need for hypothetical propositions.

adult roles. In Piaget's view, roles are systems of values and ideals which serve to control and direct the individual's behavior. Roles are adopted only in early adolescence because ideals and values are possibilities, not realities, and therefore can only be constructed by means of formal operations.

Prior to the adoption of adult roles, conceptualization was directed by the activity of concrete mental operations and the problems posed in structuring the physical environment. With the appearance of formal operations and adult roles at ages 11 to 12, these come to play the leading part in the direction of conceptualizing activity.

With social roles directing cognitive activity, the adolescents' conceptualizing becomes much more selective than it was during childhood. Unlike the relatively constant physical environment which poses similar problems for all youngsters, the diversity of social roles makes for enormous variety of concept formations among adolescents. For example, the young person who wishes to become a scientist will meet a different set of experiences—and hence conceptual problems—than will another young man whose aim is to be a mechanic. And, in general, adolescents will attain different conceptions according to the kinds of experiences and conceptual problems their particular roles provide. Accordingly, those adolescents who attain the volume conception, despite the attenuation of motivation, would have adopted roles conducive to the formation of quantity conceptions.

These two events—(a) the attenuation of interests and motivations for quantity conceptualizing due to the development of formal operations and (b) the increased selectivity, at the same age, of conceptualizing due to the adoption of diverse social roles—can explain why so many young people who spontaneously discovered the conservation of mass and weight nevertheless failed to discover the conservation of volume. It remains to determine whether this conclusion is consistent with the relations of age, sex, and IQ to volume conservation.

Age. The increase with age in the percentage of students having abstract volume conceptions is misleading because the number of students who dropped out of school also increased with age (Table 1).⁷ The groups in the higher grades were therefore a much more select sample than were the groups in the lower grades. What the relation of age and volume conservation shown in Table 1 really means is that those young people who remain in school are more likely than those who leave school to have attained an abstract conception of volume.

This finding is in agreement with the hypothesis that individual differences in social roles leads to selective conceptual learning. Students who stay in school are more likely than those who leave to have adopted professional or academic occupational roles. It seems reasonable to assume that

⁷ School records showed that the *N* for the 7th grade was average for the initial size of the groups at the older age levels. The difference between the number of 7th graders and the number of the higher grades indicates the number of dropouts at each grade level.

the motivations and interests associated with the adoption of such roles would be more likely than nonacademic or nonprofessional roles to lead individuals into situations requiring the conceptualization of quantity.

Sex. The consistently higher percentage of boys than girls having an abstract volume conception can hardly be attributed to innate differences in conceptual ability between the sexes. For one thing, the mean IQ of girls is higher than that of boys at each age level. Secondly, there is no difference between sexes with respect to the attainment of mass and weight conceptions. It seems very unlikely that there is an ability specifically for the attainment of the volume conception. Finally, Hurd (8) has shown that even when girls are far behind boys in their knowledge of physics they are almost completely able to catch up with additional training. For all these reasons differences in conceptualizing ability between sexes seems an improbable explanation for the difference in their performance.

The performance differences between boys and girls are consistent with the hypothesis presented earlier. The roles adopted by the American girl differ considerably from the roles adopted by the boy (1). Girls are supposed to be more interested in social relations than boys and to find mathematics and physics incomprehensible.⁸ Boys, on the other hand, are supposedly awkward in social situations and intuitively adept at science and mechanics. Many more boys than girls have experiences in measuring and building things, and most boys and few girls get scientific and mechanical books and toys as gifts. This difference in experiences and opportunity for developing quantity conceptions could account for the fact that more boys than girls attain an abstract conception of volume.

IQ. Individual differences in social roles could also explain the low but positive correlation of IQ with volume conservation. The IQ is neither a necessary nor a sufficient condition for the adoption during adolescence of any particular occupational role. Children with high IQs may never adopt a professional or academic ideal, while youngsters with low IQs may set their hearts on being doctors or lawyers. With the adoption of the appropriate role, even adolescents with MAs of 8 or 9 could attain the volume conception since it can be formed with the mental operations present in the average seven-year-old. One would, therefore, not expect a one-to-one correlation between volume conservation and IQ.

Nevertheless, brighter children would choose professional or academic roles more frequently than children with less ability. There would then be a tendency for bright children, more frequently than dull children, to attain the abstract conception of volume. This tendency would be reflected in a low but positive correlation of volume conservation and IQ in agreement with the correlation obtained in the present study.

In sum, the results of this investigation suggest that the appearance of formal operations and the adoption of adult roles beginning at ages 11 to 12

⁸ For example, one girl who had an IQ of 140 but failed the volume test told me her father did not believe in women getting too much education.

results in an attenuation and selectivity of inductive conceptualizing that must be taken into account before generalizing about the cognitive development of adolescents.

SUMMARY

Four hundred and sixty-nine junior and senior high school students were tested for their conceptions of mass, weight, and volume. Results showed: (a) of the students tested, 87 per cent had attained abstract conceptions of mass and weight, but only 47 per cent had attained an abstract conception of volume; (b) the percentage of students having an abstract volume conception increased significantly between the ages of 12 and 18; (c) a significantly higher percentage of boys than girls attained an abstract conception of volume; (d) there was a low but positive correlation of IQ with attainment of the volume conception.

The results of the study were interpreted within the framework of Piaget's genetic theory of intelligence. According to this theory, the operational and structural readiness for the attainment of the volume conception appears just at the onset of formal operations and the adoption of adult roles. These two events produce an attenuation of interest and opportunities for spontaneously attaining quantity conceptions whose formation is now dependent upon the particular role the young person adopts. This conclusion was found to be consistent with the relations of age, sex, and IQ to the attainment of an abstract conception of volume.

REFERENCES

1. AUSUBEL, D. *Ego development and the personality disorders*. Grune & Stratton, 1952.
2. BALDWIN, A. L. *Behavior and development in childhood*. Dryden, 1955.
3. DENNIS, W. Animistic thinking among college and university students. *Sci. Mon.*, 1953, 76, 247-250.
4. DEUTSCHE, J. M. The development of children's concepts of causal relations. *Univer. Minn. Inst. Child Welf. Monogr.*, 1937, No. 13.
5. ELKIND, D. The development of quantitative thinking: a systematic replication of Piaget's studies. *J. genet. Psychol.*, 1961, 98, 37-46.
6. ELKIND, D. Children's discovery of the conservation of mass, weight and volume: Piaget replication study II. *J. genet. Psychol.*, in press.
7. GRECO, P. & PIAGET, J. *Apprentissage et connaissance*. Paris: Presses Universitaires de France, 1959.
8. HURD, A. W. Sex differences in achievement in physical science. *J. educ. Psychol.*, 1934, 25, 70.

CHILD DEVELOPMENT

9. JERSILD, A. T., & TASCH, R. J. *Children's interests*. Bureau of Publications, Teachers College, Columbia Univer., 1949.
10. MONASH, L. Why children like or dislike school. *Understanding the Child*, 1947, 16, 67-70.
11. PIAGET, J. *The child's conception of number*. London: Kegan Paul, 1952. (First published in Switzerland, 1941.)
12. PIAGET, J., & INHELDER, B. *Le developpement des quantités chez l'enfant*. Neuchatel & Paris: Delachaux & Niestle, 1941.
13. PIAGET, J., & INHELDER, B. *The growth of logical thinking from childhood to adolescence*. Basic Books, 1958.
14. TENENBAUM, S. Attitudes of elementary school children to school, teachers and classmates. *J. appl. Psychol.*, 1944, 28, 134-141.

STUDIES OF DISABILITY: A TECHNIQUE FOR MEASUREMENT OF PSYCHOLOGICAL EFFECTS¹

ROBERT W. DOWNING, GEORGE MOED,*
and BYRON W. WIGHT

Children's Seashore House

This is the first report from a research project which is investigating social and psychological effects of disability and institutionalization on the growth and development of children. Disabilities can be classified according to severity, reversibility, duration, etc. Comparison of groups of children who differ with respect to such characteristics should ultimately reveal kinds and extents of psychological effects attributable to disability and institutionalization. Another objective of our research is to test parts of the major theories of disability. There are those who feel that the personality functioning of the disabled is best understood as reactions to being discriminated against (1, 3, 7), whereas others stress the importance of variations in body image (4). These ideas can be tested if appropriate techniques are available.

The concept of self-evaluation is crucial to an understanding of the role of body image, discrimination, or any other psychological process in disabled children. It seems reasonable that perception of one's functional

* Research Project, Children's Seashore House, Atlantic City, New Jersey.

¹ The authors are collaborators in the Research Project at Children's Seashore House, a project supported in part by a grant from the Ford Foundation and directed by Dr. George Moed. The data in this study were collected as part of the project.

The authors wish to thank their Advisory Committee members, Drs. Wilton M. Krogman, Samuel M. Wishik, and M. Erik Wright, as well as the Medical Director, Dr. Harvey N. Vandegrift and consultants, Drs. James C. Diggory, John W. Dyckman, and Seymour Feshbach, for their valuable aid and criticism.

limitations in achieving certain goals can lower self-evaluation. Equally reasonable is the idea that acceptance by the disabled of the belief that they are less worthy (because they are less capable or less attractive) than the intact can also lower self-evaluation. In either case, invidious comparisons may follow with consequent damage to the psychological well-being of disabled children and obstruction of the ordinary course of their growth and development.

This methodological report describes a semantic differential. This version of a research technique created by Osgood, Suci, and Tannenbaum (6) was constructed by us for the study of effects of disability on growth and development, self-evaluation and attitudes toward disability of disabled and intact children, and for the test of hypotheses implied by theories of disability. Reliability, factor independence, and the range of applicability of the test were investigated. Although several investigators have adapted the semantic differential technique for studies of health and disability, they have not dealt with these problems.²

Reliability, or consistency, of measurement is perhaps not so important if one is dealing with effects so powerful that they are obvious, or if hypotheses related to them are confirmed. If one is attempting to measure weaker influences, or if negative results are obtained, reliability becomes important. Before it can be concluded that an effect or hypothesized relationship does *not* obtain, it is necessary to ascertain reliability since an unreliable instrument can also yield negative results.

Factor independence must similarly be determined for several reasons. The test employs particular concepts designed specifically for use with disabled children, and the knowledge that the test consists of specific factors, rather than unknown components, is valuable in interpreting results. Moreover, the establishment of factor independence may serve as the foundation for testing predictions of circumstances under which independence will vanish and factors will be correlated. Fundamental correlation (nonindependence) of factors would preclude a definitive test of such predictions. The test to be described, for example, contains factors of Evaluation, Potency, and Activity, all of which are especially relevant to physical disability. An instance in which they might be predicted to be correlated would be that in disabled children inactivity leads to a decrease in feelings of potency and a low self-evaluation. This relation could hold notwithstanding the general independence of these factors.

Applicability involves a consideration of the limits of the test in terms of age and sex of subjects. It is our intention to develop this semantic differential not only for research but also as a clinical instrument amenable to qualitative interpretation on an individual basis. This may be accomplished if reliability, factor independence, and range of applicability are satisfactory.

² We wish to express our gratitude to Mrs. Judith Goldston and Drs. Cynthia Deutsch, Bernard Mausner, and Franklin Shontz for sharing their ideas and data with us.

METHOD

Test

This technique for measuring connotative meaning has been fully described by Osgood *et al.* (6). Our test employed nine scales³ from the three major independent factors of connotative meaning extracted by Osgood *et al.* (6). The Evaluative factor was represented by the good-bad, nice-awful, and clean-dirty scales, the Potency factor by the strong-weak, large-small, and hard-soft scales, and the Activity factor by the moving-still, hot-cold, and fast-slow scales.

Research hypotheses dealing with self-evaluations, reactions to disability and relevant persons and objects, feelings toward parents and peers, and attitudes toward cultural ideals and terms of opprobrium led to the selection of the following 15 concepts:

BOY	CRUTCH	MOTHER	OLD MAID
HOSPITAL	DOCTOR	FOOTBALL PLAYER	NURSE
ME	WHEELCHAIR	FATHER	MOVIE STAR
GIRL	SISSY	CRIPPLE	

For example, it was hypothesized that, as time since recovery from illness and discharge from the hospital increases, identification with the disabled decreases, identification here being measured as the difference in connotative meaning between ME and CRIPPLE. Both scales and concepts were chosen with regard for the age and limits of sophistication of the subjects. On this account, a five-point scale was used, quantified by the words "Very" and "Pretty." The format is illustrated below.

DOCTOR

(Very) (Pretty) (Pretty) (Very)

good : : : : bad

Procedure

Subjects were instructed⁴ to describe each of the 15 concepts by marking one of the spaces on each scale line and not to make more than one mark

³ Analysis of data from this study has since led to revision of the instructions, replacement of the hot-cold scale by a sharp-dull scale, and addition of the concept NEGRO. However, this report is based on the original version.

⁴ The following supplementary material has been deposited as Document number 6686 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C.: (a) Children's Seashore House Semantic Differential rating sheet, (b) instructions for administration of the rating sheet, (c) chance distribution of step deviation scores, (d) chance distribution of factor scores, and (e) group mean factor scores. A copy may be secured by citing the Document number and by remitting \$2.50 for photoprints, or \$1.75 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

CHILD DEVELOPMENT

on a line. They were told to use the first space if, for example, they thought DOCTOR was very good, etc. The middle (neutral) space was to be used if they thought the scale words could not be applied to the concept. Thus, each child made 135 marks ($9 \text{ scales} \times 15 \text{ concepts}$) in about 10 minutes.

Subjects

The data were obtained from a group of 41 children, ranging in age from 7 to 19 years, who had been hospitalized and immobilized for about two years but had since returned to the community in a relatively intact condition. This sequence of events occurs in Legg-Perthes disease, a disorder involving the femoral capital epiphysis of children (5). The group is especially interesting because they have moved from a situation of immobilization and disability to one of mobility and intactness. What are their identifications, and do their attitudes toward the disabled differ from either intact children who have never been disabled or from those of permanently disabled children? Still another interesting question concerning this group is that, if disability entails value loss (2), is recovery accompanied by value gain?

Table 1 presents data describing the subjects. The ratio of boys to girls (29:12) is a reflection of the differential frequencies with which they are affected with the disease. In addition to data on sex differences, age differences were analyzed. The total group was divided into a "Young" and an "Old" group by splitting the sample at the median and discarding the middle case. This split placed roughly half of the boys and half of the girls in each age group. Chi squares calculated for differences between the Boy-

TABLE 1
SEMANTIC DIFFERENTIAL SAMPLE CHARACTERISTICS
(All figures in months)

	Total N=41	Boys N=29	Girls N=12	Young N=20	Old N=20
<i>Age at Testing</i>					
Mean	165	162	171	133	196
Range	87-227	87-217	112-227	87-165	173-227
<i>Length of Hospitalization</i>					
Mean	24	24	24	21	27
Range	2-46	2-46	19-41	7-37	2-46
<i>Time since Discharge</i>					
Mean	63	59	71	37	86
Range	8-107	8-107	14-106	8-77	37-106

DOWNING, MOED, and WIGHT

Girl and Young-Old subgroups in length of hospitalization were not significant. Whereas the Old group had, of course, a longer time since discharge, there was no such difference between the Boys and Girls.

Table 2 shows prorated Verbal Intelligence Quotients. Tests of differences between subgroups were not significant. The total group was slightly, but significantly, lower than the groups used in establishing norms for the Wechsler Intelligence Scale for Children and the Wechsler Adult Intelligence Scale.

TABLE 2
VERBAL INTELLIGENCE QUOTIENTS*

	Total	Boys	Girls	Young	Old
Mean	95.1	95.7	93.5	94.1	95.1
SD	14.5	12.9	17.8	15.7	12.9
Range	63-124	63-120	66-124	63-120	66-124

* Prorated Verbal Intelligence Quotients based on the Wechsler Intelligence Scale for Children and the Wechsler Adult Intelligence Scale Information, Comprehension, Similarities, and Vocabulary subtests.

Measures of Performance

Reliability. The reliability measure, similar to one described by Osgood *et al.* (6), was developed to answer the question: Are the subjects consistent in using the three scales which represent each factor?

The numbers 2, 1, 0, -1, and -2 were assigned to the points on each scale from left to right. Thus, for each concept, each subject had three scores for each factor. If a subject marked a concept at exactly the same point on the three scales representing a factor then his within-factor reliability was maximal, and the greater the scatter on these scales the lower the reliability.

For each concept the three scale markings of a factor were treated as combinations of three pairs of scores. A difference score (always considered positive) was obtained for each of the pairs. The sum of these differences is the "step deviation," or reliability score, for one subject rating one concept on one factor. If, for example, a subject obtained the following scores on the Evaluative factor for the concept ME

good-bad 2
clean-dirty 1
nice-awful 0

then his step deviation score was $(2 - 1) + (2 - 0) + (1 - 0) = 4$. With this method of scoring the only possible step deviation scores are 0, 2, 4, 6, and 8. Zero step deviations come from three identical scale ratings

while scores of 2 require two identical ratings with a third only one position away.

It seems reasonable to assume that scores greater than 2 step deviations reflect unreliability of the subject, or the scales, or a significant scale-concept interaction. Since only two kinds of combinations yield scores of 0 or 2, the reliability of these scores is unambiguous and their occurrence by chance alone is small. However, many combinations can generate scores of 4, 6, and 8 so that the meaning of these scores is not known to us and they can more readily be attributable to chance. Setting the reliability criterion at scores of 0 and 2 means that no subject's ratings can be counted reliable if they lie on both sides of the neutral space, the implication being that such ratings reflect a conflict of feeling about a concept. In assessing reliability, the frequencies of obtained step deviation scores of 0 and 2 were compared with those expected by chance. Chance frequencies were determined by generating the distribution of all possible combinations of three scale scores and counting the numbers of ways of obtaining each of the possible step deviation scores (0, 2, 4, 6, and 8). Scores of 0 and 2 are expected, on the average, 23 per cent of the time by chance alone.⁴

Step deviation scores of 0 and 2 were summed for each concept on each factor separately, yielding three distributions of reliability scores in which the 15 concepts were the *N*. If the factors are reliable, then frequencies of 0 and 2 scores in each one of these distributions should be significantly large. If reliabilities of the factors are independent, then correlations between pairs of them should be nonsignificant. If the instrument has a wide applicability, then tests of differences, on this index, between either age or sex subgroups should be nonsignificant.

Factor scores. Just as the reliability score represents variability of the scale scores, factor scores represent their central tendency. For each concept a subject's scores on the three scales representing a factor were averaged to obtain his factor score. Following the illustration above where a subject marked the Evaluative scales 2, 1, and 0 for the concept ME, his factor score was $(2 + 1 + 0) \div 3 = 1$. Factor scores derived from the five-point scale can range from +2 to -2. Scores as large as 2 can be obtained only from three identical scale scores of 2. The probabilities of obtaining them by chance alone are much smaller than the probabilities of obtaining smaller factor scores (like 1 or 0) since these may be arrived at through many different combinations of three scale scores. When it was desired to compare obtained factor scores with those expected by chance, these differential probabilities proved valuable since they are finite, exact, and calculated.⁴

The psychological significance of these factors lies in their relevance for describing the self-perceptions of disabled children and their ratings of important persons and objects. In the test they are used to index the amounts and directions of evaluation, potency, and activity attributed to concepts. One might hypothesize that children who are severely disabled would attribute a higher evaluation to concepts like DOCTOR and NURSE (upon

whom they are quite dependent) than children suffering less incapacitation. At a more descriptive level it might be interesting to compare disabled children's self-ratings of potency with those of intact children.

Group mean factor scores were derived for each concept on each factor separately, yielding three distributions of group factor scores in which the 15 concepts were the *N*. If the factors are independent, then correlations between pairs of these distributions should be nonsignificant. If the instrument has wide applicability, then tests of differences, on this measure, between either age or sex subgroups should be nonsignificant.

Meaning. Geometrically, the E, P, and A factor scores are the coordinates, in a three dimensional semantic space, of a point representing a concept for a subject. The point where these dimensions bisect each other and form perpendicular axes represents zero, or the origin, in the space. Meaning, as a point in space, can be composed of positive as well as negative factor scores in addition to magnitudes of these scores. Thus, it is possible to characterize concepts and to compare characterizations derived from responses of individuals and groups. The meanings attributed to concepts concerning disability can be better understood by measuring their relative

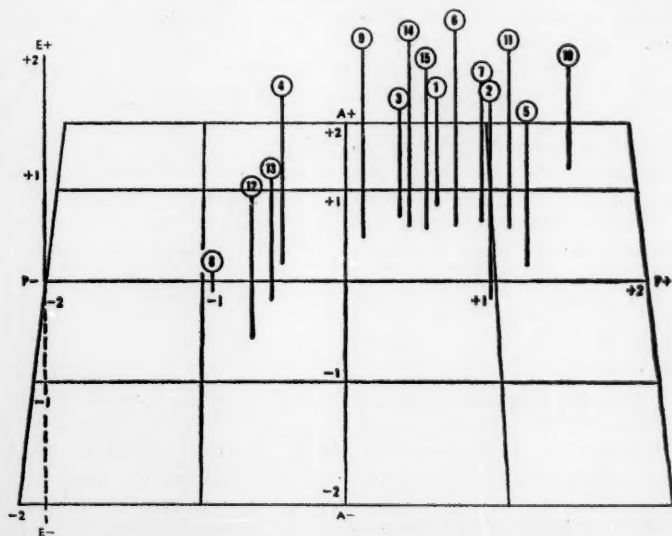


FIGURE 1—Semantic space for former Legg Perthes patients. Concepts: (1) boy, (2) hospital, (3) me, (4) girl, (5) crutch, (6) doctor, (7) wheelchair, (8) sissy, (9) mother, (10) football player, (11) father, (12) cripple, (13) old maid, (14) nurse, (15) movie star.

proximity to concepts which have generally stereotyped connotations, such as the ideals and terms of opprobrium included in the test. The concept FOOTBALL PLAYER should be characterized positively while Sissy should be characterized negatively. Will CRIPPLE be characterized similarly to FOOTBALL PLAYER, Sissy, or neither? Discernment of relations among the characterizations of concepts could be used in personality investigation and in the test of theories of disability.

One approach to the analysis of meaning was to construct a three dimensional model of the semantic space and mark the locations of the concepts. This model (see Figure 1) facilitated the interpretation of each concept's characterization as well as the relations between characterizations. The limits of applicability of the test for the age and sex subgroups were explored by testing for the significance of differences in meaning between subgroups on each concept. Finally, the differences between obtained meanings for each concept and those expected by chance were calculated to test the appropriateness of the concepts.⁴

Distance. Osgood *et al.* (6) have detailed the method of investigating distances between concepts. Just as any concept can be compared with a zero point, or origin, in a semantic space, so can distances between concepts be compared. As expressed in the formula below, distance is defined as the square root of the sum of the squared differences between factor scores for any two concepts.

$$D_{12} = \sqrt{(E_1 - E_2)^2 + (P_1 - P_2)^2 + (A_1 - A_2)^2}$$

Self-perceptions can be interpreted in relation to their proximity to other concepts even apart from the locations of these concepts in space. Of obvious interest are such comparisons as the distance between ME and CRIPPLE with ME and BOY (or GIRL). The distance approach is an avenue for the study of identification, self-evaluation, and some components of personality functioning of the disabled.

RESULTS

Reliability

Analysis of reliability scores of 0 and 2 by concepts over factors showed that all concepts exceeded chance frequencies of 23 per cent on the Evaluative factor in all groups. On the Potency factor all concepts exceeded chance frequencies in all groups except MOTHER in the Girls group. The only instances of concepts failing to exceed chance on the Activity factor were HOSPITAL in the Girls and Young groups and WHEELCHAIR in the Old group. The results also showed that reliable responses were fairly evenly distributed among the concepts.

Chi square analysis of reliability scores of 0 and 2 by subjects over factors showed that significant numbers of subjects exceeded the 0 and 2 criterion

TABLE 3

NUMBERS OF SUBJECTS WHO EXCEEDED THE CHANCE FREQUENCY
OF 0 AND 2 RELIABILITY SCORES

Factor	Total	G R O U P			
		Boys	Girls	Young	Old
E	41	29	12	20	20
P	39	28	11	18	20
A	31	19	12	13	17

on each factor (see Table 3). Table 4 shows 0 and 2 scores as percentages of all scores; with the exception of the Young group on the Activity factor all percentages were significantly large.

Analysis by concepts, subjects, or responses pointed to the Activity factor as the least reliable, and examination of this factor showed that the lower reliability was attributable to a large number of neutral responses on the hot-cold scale (see Table 5). Evidently, this scale was inappropriate for some concepts and subjects. (In order to achieve greater reliability, the hot-cold scale has since been replaced by a sharp-dull scale in a revised form of the test.)

Although the Evaluative, Potency, and Activity factors all showed reliability significantly greater than chance, they differed considerably on this measure ranging from high to low in the order mentioned. In Table 6 the between-factor rank order correlations of reliability scores for all groups were not significant, reflecting factor independence as measured by 0 and 2 scores.

Table 7 provides data concerning reliability score differences in the age and sex subgroups. The Wilcoxon matched-pairs signed-ranks test was used to evaluate these differences. On the Potency and Activity factors the Old

TABLE 4

0 AND 2 RELIABILITY SCORES PRODUCED BY EACH GROUP ON EACH FACTOR
(All figures expressed in percentages)

Factor	Total	G R O U P			
		Boys	Girls	Young	Old
E	80	81	79	77	83
P	53	53	52	46	59
A	42	42	43	35	47

CHILD DEVELOPMENT

TABLE 5
NEUTRAL RESPONSES BY SCALES AND FACTORS

<i>Scales</i>	<i>N</i>	<i>%</i>
good-bad	43	3
nice-awful	63	4
clean-dirty	46	3
Σ Evaluative	152	10
strong-weak	66	5
hard-soft	208	15
large-small	212	15
Σ Potency	486	35
moving-still	160	11
hot-cold	450	32
fast-slow	172	12
Σ Activity	782	55
Total	1420	100

group was clearly more reliable than the Young group. This higher reliability, not attributable to sample size or IQ measures, probably reflects the greater cognitive and affective differentiation of the older children. As can be seen from Tables 3 and 4, lower reliability patterns came from the Young group on all factors but especially from the Activity factor. Interpretation of responses from children much younger than this sample or from the Activity factor alone should be made with caution.

Thus far, the problems of reliability, factor independence, and general applicability of the test have been examined by using step deviation scores of 0 and 2. The results, by and large, confirmed the independence of the factors and established the reliability and applicability of the test with certain limitations relating to age and to the Activity factor.

TABLE 6
BETWEEN-FACTOR CORRELATIONS OF RELIABILITY

<i>Factor Comparisons</i>	<i>Total</i>	<i>Boys</i>	<i>Girls</i>	<i>Young</i>	<i>Old</i>
E - P	-.20	-.14	-.19	-.10	-.01
E - A23	.36	.07	-.18	.33
P - A	-.21	-.11	-.34	-.16	-.19

DOWNING, MOED, and WIGHT

TABLE 7
BETWEEN-GROUP DIFFERENCES IN RELIABILITY

Factor	SMALLER SUM OF RANKS	
	Boys-Girls	Young-Old
E	40.0	16.5
P	58.5	11.0*
A	49.5	16.5*

* $p < .05$.

Factor, Meaning, and Distance Scores

Because of the interrelations of factor, meaning, and distance scores, they are treated together in the results which follow. A preliminary part of the analysis was undertaken to determine appropriateness of the concepts by comparing the obtained distributions of factor scores for each concept with the distribution expected by chance. Forty-five comparisons (15 concepts on three factors) were made using the Kolmogorov-Smirnov one-sample test (see Table 8). There were 39 factor distributions different from

TABLE 8
MAXIMUM PERCENTAGE DEVIATION FROM CHANCE OF OBTAINED
FACTOR SCORE DISTRIBUTIONS

Concept	Evaluation	Potency	Activity
Boy41*	.43*	.52*
Hospital85*	.45*	.16
Me62*	.26*	.43*
Girl77*	.29*	.20
Crutch67*	.67*	.28*
Doctor85*	.40*	.38*
Wheel Chair67*	.52*	.40*
Sissy09	.52*	.20
Mother82*	.13	.30*
Football Player43*	.72*	.70*
Father79*	.58*	.38*
Cripple62*	.36*	.36*
Old Maid57*	.31*	.18
Nurse84*	.33*	.40*
Movie Star74*	.33*	.40*

* $p < .05$.

CHILD DEVELOPMENT

chance at the .05 level, indicating a high degree of agreement among subjects as to the connotative meaning of these concepts. The six nonsignificant results were: Sissy on E; MOTHER on P; HOSPITAL, GIRL, Sissy, and OLD MAID on A. The possibility that these results might be due to offsetting directions of meaning attributed to the concepts by the subgroups was examined and rejected except for the one instance of HOSPITAL on the A factor. The extent to which these distributions are due to ambivalence of meaning induced in individual subjects by these concepts requires further study.

Further analysis of factor scores was undertaken to determine the agreement of the subgroups on connotative meanings attributed to the concepts. Differences in factor score distributions were tested for each concept with the Kolmogorov-Smirnov two-sample test (*see* Table 9), and chi squares for the three factors were summed inasmuch as reliabilities (errors of measurement) of the factors have been shown to be independent.

In only one instance was the *combined* chi square significant at the .05 level, indicating a significant distance between subgroups. Boys and Girls disagreed on the characterization of the concept GIRL with Girls attributing greater potency and activity to this concept. There were several additional instances where significant differences were found on *individual* factors. Girls rated CRUTCH as more active. The Old group had a more positive evaluation of WHEELCHAIR and OLD MAID and attributed greater potency to ME and more activity to HOSPITAL than the Young group.

TABLE 9
BETWEEN-GROUP DIFFERENCES IN MEANING

Concept	COMBINED CHI SQUARES	
	Boy-Girl	Young-Old
Boy	3.00	8.80
Hospital	5.25	10.10
Me	4.29	10.10
Girl	15.00*	6.70
Crutch	10.02	3.80
Doctor	8.04	2.20
Wheel Chair	5.07	9.30
Sissy	4.62	11.00
Mother	7.35	5.70
Football Player	2.47	4.90
Father	7.57	1.70
Cripple	2.49	4.30
Old Maid	5.40	10.90
Nurse	6.91	5.40
Movie Star	6.85	2.20

* $p < .05$.

DOWNING, MOED, and WIGHT

Further evidence for the applicability of the test may be seen in the correlations of factor scores in Table 10. Subgroups generated similar orders of concepts. However, the Girls rated more positively than the Boys on all factors and most concepts. This may reflect the operation of a cultural norm which permits greater expression of affect on the part of females. The small

TABLE 10
BETWEEN-GROUP CORRELATIONS OF FACTOR SCORES

<i>Factor</i>	<i>Boys-Girls</i>	<i>Young-Old</i>
E91*	.85*
P98*	.93*
A91*	.92*

* $p < .05$.

number of significant differences and the high correlations between subgroups in the rating of concepts means that, by and large, subgroups can be treated together in the testing of hypotheses and that separate norms are not required for interpretation of results. General applicability of the test was thus supported.

Unique characteristics of the subjects and the concepts demanded a test of factor independence. Rank order correlations between group mean factor scores were calculated (*see* Table 11). The E factor was independent in all comparisons. Although only one correlation was significant, those in the P-A comparisons were higher than the others. In general, it may be said that factor independence was sustained with these peculiar characteristics of concepts and subjects.

TABLE 11
FACTOR INTERCORRELATIONS

<i>Factor Comparisons</i>	<i>Total</i>	<i>Boys</i>	<i>Girls</i>	<i>Young</i>	<i>Old</i>
E - P23	.23	.16	.29	.15
E - A	-.02	-.17	.20	-.06	-.07
P - A49	.33	.48	.37	.54*

* $p < .05$.

Figure 1 shows the locations of the concepts as characterized by the total group. None of the concepts had negative E scores, although one was close to zero. Perhaps the children reflected a cultural prohibition against evaluating objects or persons negatively in the presence of adults (the experimenters). The most negatively characterized concept was Sissy, and

CHILD DEVELOPMENT

the most positively characterized concept was FOOTBALL PLAYER as expected. The psychological meaning of concepts can be examined by measuring their proximity to such reference points.

Concepts which were closest to Sissy were OLD MAID, CRIPPLE, and GIRL. The proximity of OLD MAID is expected since it was included as a control concept for Sissy. The proximity of GIRL and CRIPPLE to such terms of opprobrium is predicted from a general theory of discrimination against the disabled, similar to race or sex discrimination (7). Subjects characterized these concepts as lacking potency and activity, ingredients which may be embedded in the stereotypes of these concepts. Although the Girls joined in the stereotype of GIRL, they were not as negative as the Boys, as noted above. Conversely, concepts like FATHER and DOCTOR were found in the vicinity of FOOTBALL PLAYER in the most positive octant of the space.

These patterns of group factor and meaning scores have been presented as examples of the potentialities of these measures for research on disability. New data have since been obtained from disabled and intact children which will enable us to evaluate the responses of the original group of formerly disabled children.

SUMMARY AND CONCLUSIONS

This methodological paper has described a version of a semantic differential developed for investigating the psychological effects of disability and institutionalization on the growth and development of children. The subjects had all been hospitalized and immobilized for Legg-Perthes disease, returned to the community when they recovered, and tested when they were recalled for follow-up study. In addition to describing the subjects, procedure, and measures of performance, data have been presented concerning reliability, independence of the Evaluative, Potency, and Activity factors, and general applicability of the test for age and sex subgroups.

The children were more consistent in rating the Evaluative scales than scales of Potency and Activity. This observation raises the possibility that independence of factors (in this context) may be attributable to different levels of task difficulty for the subjects, an idea which merits investigation. It was also observed that reliability increases with age, which coincides with the view that there is greater cognitive and affective differentiation in older children. A further finding which has developmental implications is that girls rated concepts more positively than boys.

The authors feel that the following may be concluded:

1. The factors (in all tabulations) were significantly more reliable than chance, and reliable responses were fairly evenly distributed among the concepts indicating that the test is reliable and the choice of concepts is satisfactory.

2. Between-factor correlations of both reliability scores and factor scores, in general, sustained the independence of the factors with these subjects and concepts.

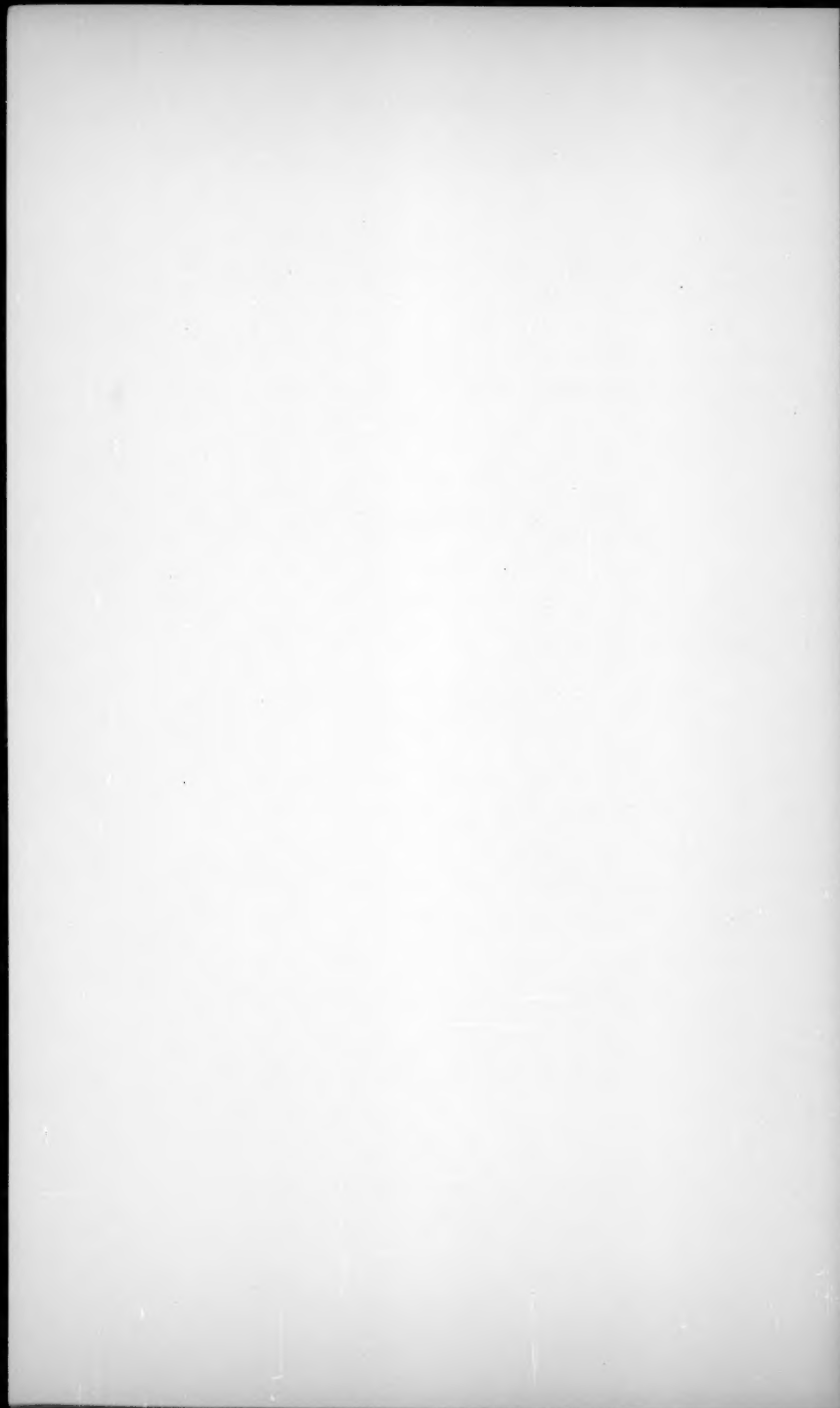
3. Meaningfulness of the concepts was confirmed by the fact that subjects responded to them systematically, rather than randomly.

4. Since there was only one significant difference between subgroups in the meaning attributed to concepts and they generated very similar orders of concepts, the test appears to have rather wide applicability.

5. Preliminary examination of group meaning data suggested that this semantic differential may be used to assess the role of disability in personality development and to investigate discriminatory attitudes toward the physically disabled.

REFERENCES

1. CRUICKSHANK, W. M. (Ed.) *Psychology of exceptional children and youth*. Prentice-Hall, 1955.
2. DEMBO, T., LEVITON, G., & WRIGHT, B. Acceptance of loss—amputations. In J. F. Garrett (Ed.), *Psychological aspects of disability*. U. S. Department of Health, Education, and Welfare, Office of Vocational Rehabilitation, Rehabilitation Service Series No. 210, 1952. Pp. 80-96.
3. DEMBO, T., LEVITON, G., & WRIGHT, B. Adjustment to misfortune—a problem of social-psychological rehabilitation. *Artificial Limbs*, 1956, 3, 4-62.
4. FISHER, S., & CLEVELAND, S. E. *Body image and personality*. Van Nostrand, 1958.
5. GOFF, C. W. *Legg-Calve-Perthes syndrome*. C. C. Thomas, 1954.
6. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Univer. of Illinois Press, 1957.
7. WRIGHT, B. A. *Physical disability—a psychological approach*. Harper, 1960.



RESPONSE SETS AND MANIFEST ANXIETY SCORES IN A RETARDED POPULATION

CHARLES G. MATTHEWS*

Fort Wayne State School

and LEON H. LEVY

Indiana University

The prominence of the construct of anxiety in the thinking of psychologists demands that any instrument purporting to yield a measure of it be subjected to the closest possible scrutiny. The present study has as its purpose, therefore, an examination of the applicability of the Children's Form of the Manifest Anxiety Scale (CMAS) (1) for use with mentally retarded population. Since the CMAS was standardized on normal children in the fourth, fifth, and sixth grades and since it has been shown to be related to intelligence and school achievement (2, 6), one might question the meaning of scores for mental defectives on this scale. In the present study this problem was investigated in several ways. The first was by an examination of the test performance per se of retardates, i.e., their group mean and standard deviation, test-retest reliability, and item analysis data; the second, more directly aimed at the construct validity of the CMAS when used with mental defectives, involved the correlation of CMAS scores with performance on a digit span recall test based on the hypothesis that anxiety interferes with the abilities tapped by the digit span test. That is, from a variety of different theoretical positions, anxiety has been conceptualized in such a way that high anxiety would be expected to interfere with learning and performance. Thus, if the CMAS is a valid measure of the construct of anxiety, scores on it should correlate negatively with a digit span recall test. Although the evidence for the validity of this hypothesis is far from univocal (4, 5, 7, 8), finding a significant negative correlation could be taken as supporting evi-

* Department of Psychological Services, Fort Wayne State School, Fort Wayne, Indiana.
Child Developm., 1961, 32, 577-584.

CHILD DEVELOPMENT

dence for the construct validity of the CMAS when used with a retarded population.

Still related to the meaning of CMAS scores for a retarded population, but possibly having wider implications, is the question of the contribution of response sets. Because items only contribute to the anxiety score if answered "Yes," it seems reasonable to hypothesize that acquiescent or assent tendencies contribute significantly to the score obtained, this, all the more so, for a retarded population, because of the common observation of the greater acquiescence and persuasability of this group. Since, with the exception of two items, the 11 items constituting the Lie Scale of the CMAS contribute to the L Score when answered "Yes," we might expect acquiescence to affect this score as well. Thus, the third approach to the meaning of CMAS scores for retardates involved the correlation of these scores with a specially constructed instrument to measure assent and dissent tendencies. From the standpoint of the response set hypothesis, the predictions were made that both the anxiety score and the lie score would correlate positively with the measure of assent tendencies and negatively with the measure of dissent tendencies.

METHOD

Subjects

Thirty mildly retarded male residents of the Fort Wayne State School were used in this investigation. Etiological background in the experimental group was heterogeneous, with diagnostic labels of "familial" and "congenital cerebral maldevelopment—nonspecific" contributing the bulk of the cases. Patients having gross visual, auditory, or motor handicaps were excluded from the sample, as were residents with histories of psychotic episodes or other marked affective or ideational disturbances.

Subjects ranged in age from 18 to 30 years. IQ scores ranged from 48 to 72, with scores based upon 24 Full Scale WAIS tests and six Stanford-Binet (L) results obtained within two years of the time of this investigation. While it is not strictly defensible to combine scores from different psychometric tests to yield a group mean, it was felt that such a procedure would serve to describe in a general way the intellectual characteristics of the group (IQ mean = 59.5; SD = 6.7).

Procedure

Subjects were seen individually, with uniform experimental setting and procedure for each subject. Immediately prior to test administration, each subject was queried briefly about cottage placement and about recreational interests and activities in an effort to convey an attitude of informality and acceptance to the subject. Following this brief interview, the CMAS was administered to the subject. Because of the nature of the experimental population, each item in the CMAS was read to the subject and his responses recorded. CMAS instructions were given as follows:

I have a number of sentences or statements on this paper that I am going to read to you. I want you to listen carefully to each one I read, and to tell me whether the sentence says something that is true about you or not true. If the sentence I read says something that describes you or the way you feel, say "True"; if the sentence is not true or does not describe you, say "Not True."

Upon completion of the CMAS, the Digit Span subtest from the Wechsler-Bellevue I was administered to each subject, and highest number of digits recalled was recorded. In order to obtain a measure of response stability on the CMAS, the scale was re-administered to each subject two weeks after the first administration. The departure in administration of the CMAS in the present study from that used in its standardization was required for obvious reasons, and its significance in accounting for the subsequent difference found cannot be gainsaid without further empirical study. On the other hand, it might be contended that if the CMAS were to be used with a retarded population it could not be administered in any other way and so must be evaluated on these grounds.

In order to investigate possible relations between response set, CMAS scores, and Digit Span performances, the Assent-Dissent Response Set Scale (ADRSS) was developed by the authors and administered to each subject immediately following the second administration of the CMAS.

The ADRSS consists of 16 statements which are read to the subject by the examiner. The subject is requested to answer "True" or "Not True" to each statement. Eight pairs of contradictory statements are used in the ADRSS, with the statements in each pair sequentially interrupted by statements from other pairs used in the test. For example, statement 1 is: "My left arm is a little longer than my right arm"; its contradictory statement is number 10: "My arms are both the same length." Similarly, statement 3 is: "I like watching TV more than playing sports"; while statement 11 is: "I like sports better than TV."

Scoring is based upon the number of contradictory pairs in which both statements are answered in a particular direction. Since the subject can answer either "True" or "Not True," the ADRSS yields both an assent response set score and a dissent response set score. Assent response set scores may range from 0 to 8 and are based upon the number of contradictory pairs in which both statements are answered "True." Dissent response set scores may range from 0 to 8 and are based upon the number of contradictory pairs in which both statements are answered "Not True."

RESULTS

Test Performance

Table 1 presents the test and retest means and standard deviations for the anxiety and lie scores as well as their test-retest reliabilities. For comparison purposes, similar data are presented for the 73 sixth grade boys of

CHILD DEVELOPMENT

TABLE I

TEST PERFORMANCE OF RETARDATES AND SIXTH GRADE BOYS

	TEST (N=30)		RETEST (N=30)			6TH GRADE BOYS (N=73)		
	Mean	SD	Mean	SD	r	Mean	SD	r
Anxiety Score	22.2	8.7	19.0	9.4	.84	16.6	7.4	.82
Lie Score	5.1	2.3	4.4	2.7	.86	1.8	1.6	.53

the standardization group (1), the group probably closest to the present one in mental age. Here it will be noted that, for both the anxiety and lie scores, the means and standard deviations of the retarded group are larger. Comparing the first test mean anxiety and lie scores of the retardates with those of the sixth graders by means of the *t* test revealed differences significant at the .01 and .001 levels, respectively. The variances for the two groups for these two scores, when compared by means of the *F* ratio, were found to be significantly different ($p < .05$) in the case of the lie score, thereby making the *t* test for the lie score means only somewhat suspect as to meaning. Comparing test-retest reliabilities for the two groups—a comparison not entirely justified since the test-retest interval for the sixth grade boys was only one week—revealed no difference for the anxiety score, but a difference significant at the .001 level for the lie score, with the retardates showing the greater stability.

Finally, the correlation between first test mean anxiety and lie scores for the retardates was $-.04$. This compares favorably with the corresponding correlation of $-.10$ obtained for the standardization group, suggesting that in at least one aspect of their performance the two groups do not differ significantly.

To get some indication of population item-response characteristics, an item analysis was performed on the anxiety items, counting the number of "True" responses for each item for the entire group. In Table 2 are presented those items answered "True" by 80 per cent or more, or by 20 per cent or less, of the total group.

The items answered "True" by 80 per cent or more of the group, it will be noted, represent concerns which might be expected and considered appropriate for a person removed from his home and placed in an institution (items 8 and 46) and of limited ability (items 9, 29, and 44). Furthermore, of 11 items having some reference to ability, achievement, or attention span, nine were answered "True" by 50 per cent or more of the group. For these same items, Hafner and Kaplan (3) found only four on which 50 per cent or more of their 62 normal fifth grade boys answered "Yes." These data, therefore, strongly suggest that the item content of the CMAS is such as to discriminate between children of differing ability levels and possibly

TABLE 2

CMAS ANXIETY ITEMS ANSWERED "TRUE" BY 80 PER CENT OR MORE, OR BY 20 PER CENT OR LESS OF THE TOTAL GROUP

Item Number	Content	Per Cent Answering "True"
29	I worry about doing the right things	93
8	I wish I could be very far from here	90
46	I often worry about what could happen to my parents	90
44	I often do things I wish I had never done	83
9	Others seem to do things easier than I can	80
19	Often I have trouble getting my breath	20
23	I have to go to the toilet more than most people	20
26	I have trouble swallowing	03
39	I am afraid of the dark	00

also between institutionalized and noninstitutionalized children. The significant difference in mean anxiety scores, already noted, might further be taken as confirming this.

Of the four items answered "True" by 20 per cent or less of the retardates, three are concerned with somatic functioning. For these same items, Hafner and Kaplan found frequencies of 16, 7 and 5 per cent, suggesting that, if somatic complaints are highly correlated with anxiety, the groups may not be as different as their CMAS scores would suggest. More simply, perhaps, the similarity between normals and retardates in responses to the somatic complaint items may indicate that this is an area which is little affected by either intellectual ability or institutionalization.

Evidence for Construct Validity

The correlation between the anxiety score and digit span recall was .00. The corresponding correlation for the lie score was -.04. Thus, to the extent that the hypothesis linking anxiety with interference with performance on such tasks as digit span is valid, the present findings must be counted as negative evidence for the construct validity of the CMAS when used with retardates.

Evidence for Response Sets

The mean assent score on the ADRSS was 1.5 with a range of 0 to 6, while the mean dissent score was .77 with a range of 0 to 3. The correlation between assent and dissent scores was -.40, significant at the .02 level by a one-tailed test, justified in this instance by the logic of the scale which demands that the correlation be negative for the scale to have the meaning imputed to it. The correlations of assent and dissent scores with the anxiety

CHILD DEVELOPMENT

score were $+.31$ and $-.37$, respectively, both significant beyond the .05 level by a one-tailed criterion, thus supporting the response set hypothesis in connection with the anxiety score. The corresponding correlations with the lie score were $+.31$ and $+.33$, the former significant at the .05 level by a one-tailed criterion and the latter correlation being in the wrong direction. Therefore, the response set hypothesis is only partially supported in the case of the lie score.

Because it might be contended that the ADRSS correlations with the CMAS merely reflect a relation between assent and dissent tendencies and anxiety, the two ADRSS scores were correlated with digit span recall. If this reasoning were correct, assent would correlate negatively with digit span and dissent positively. The obtained correlations were $-.34$ and $+.09$, respectively, both in the proper direction for support of the above reasoning but with only the former correlation being significant at the .05 level by a one-tailed criterion. However, this finding and the reasoning behind it loses its cogency when it is remembered that the CMAS itself failed to correlate at all with digit span recall. Therefore, unless we are willing to conceive of the assent score as a better indicant of anxiety than the CMAS, there is no reason to believe that the correlations between ADRSS scores and CMAS scores represent anything other than the contribution of response sets to test score variance.

DISCUSSION

While CMAS scores of mental defectives appear to be as stable as those of normals, the results of this study point fairly conclusively to the inapplicability of the test as a measure of anxiety in mental defectives. Put more conservatively perhaps, whatever it is that the CMAS score represents in a normal population, it is evident that it represents something different for a retarded one. Not only would the CMAS be likely to portray retardates as more anxious than normals, but also as less truthful. Whether either of these is correct or not cannot be ascertained by an instrument composed of items such as those contained in the CMAS which are so obviously situationally and intellectually biased or by an instrument as susceptible to response sets as the CMAS appears to be.

The absence of correlation between digit span recall and CMAS cannot be considered too damaging and may be considered as reflecting as much upon the validity of the theory leading to the prediction that such a correlation would exist as it does upon the construct validity of the test itself.

On the other hand, the support found for the response set hypothesis in the present study points beyond the problem of the meaning of CMAS scores for retardates. For recent research demonstrates that response sets may be readily found among normals as well and that a test such as the CMAS keyed only in one direction should be considered suspect on a priori grounds alone. By constructing a separate scale to measure response sets in

the present study, rather than following the usual practice of simply reversing items in the test under study, it was believed that we would be tapping a more general disposition. Because of this, it would seem that the correlations obtained between ADRSS and CMAS scores are all the more significant, for these suggest that the response sets measured are not test specific. It is quite likely that, had the usual procedure of reversing items been followed or had the ADRSS been somewhat longer, thereby permitting a greater range of talent to be manifested, the correlations obtained would have been somewhat larger.

Although evidence has been presented for the operation of response sets in the CMAS, it is evident from test performance alone that they do not account for all of the variance. For one thing, if they did, the correlation between the anxiety and lie scores would be much higher than it is. For another, the item analysis would have revealed greater skewing of items in terms of frequency of "True" answers. Thus, of the 42 anxiety items, 26 were answered "True" by 50 per cent or more of the group; of the 11 lie items, six were answered "True" by 50 per cent or more of the group, thereby indicating in both instances only some skewing in the direction expected if an assent response set were operating.

Thus, it may be concluded that, for mental defectives, response sets account for some of the variance in the CMAS, but that a major portion of it is probably also determined by such factors as intellectual status and institutionalization. Whether any significant portion of CMAS score variance remains which may be meaningfully defined as anxiety remains uncertain.

SUMMARY

In an attempt to assess the applicability of the Children's Form of the Manifest Anxiety Scale to a mentally retarded population, the test was administered to 30 institutionalized, mentally retarded males on a test-retest basis together with a digit span test and a specially constructed response set scale. CMAS scores were found to be stable over time, but significantly higher than those for the sixth grade boys used in its standardization. No correlation was found between CMAS and digit span, but significant correlations were found with the response set scale. On the basis of these data, as well as an item analysis, it was concluded that the scores of retardates on the CMAS are strongly influenced by response sets, as well as situational and capacity variables, and that the test, therefore, is not appropriate for a retarded, institutionalized population.

REFERENCES

1. CASTANEDA, A., McCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.

CHILD DEVELOPMENT

2. HAFNER, A. J., & KAPLAN, A. M. Children's manifest anxiety and intelligence. *Child Developm.*, 1959, 30, 269-271.
3. HAFNER, A. J., & KAPLAN, A. M. An item analysis of the children's manifest anxiety scale. *Child Developm.*, 1959, 30, 481-488.
4. KELLER, J. E. The relationship of auditory memory span to learning ability in high grade mentally retarded boys. *Amer. J. ment. Def.*, 1957, 61, 574-580.
5. LEWINSKI, R. J. The psychometric pattern: I. Anxiety neurosis. *J. clin. Psychol.*, 1945, 1, 214-221.
6. McCANDLESS, B. R., & CASTANEDA, A. Anxiety in children, school achievement and intelligence. *Child Developm.*, 1956, 27, 379-382.
7. MOLDOWSKY, S., & MOLDOWSKY, P. Digit span as an anxiety indicator. *J. consult. Psychol.*, 1952, 16, 115-118.
8. SHOEN, E. J. The Wechsler Bellevue in the detection of anxiety; a test of the Raskis-Welsh hypothesis. *J. consult. Psychol.*, 1950, 14, 40-45.

RELATIONS BETWEEN NURSERY SCHOOL
ATTENDANCE AND TEACHERS' RATINGS OF SOME
ASPECTS OF CHILDREN'S ADJUSTMENT
IN KINDERGARTEN¹

ANN WILSON BROWN *and* RAYMOND G. HUNT*

Washington University, St. Louis

Perhaps one of the more significant aspects of change in western culture has to do with the increasingly prominent role played by educational institutions in child development. Progressing from relatively exclusive concern with fostering the literacy of its charges, the school has assumed primary responsibility for vocational preparation and, in contemporary concepts, has evolved as one of the salient cultural agents in the general socialization process.

Accompanying these shifts in the focus of educational concern has been an equally notable extension in time of institutional responsibility. On the one hand, secondary and higher educational activities have extended farther into young adulthood (and even beyond), while at the same time a "downward" extension of educational responsibility has steadily formed in the establishment of the kindergarten and other "preschool" programs.

All of this can undoubtedly be correlated with generalized mutations in western culture and particularly with changing concepts of the status and role of the family (11). In any event, it seems quite clear that for some time to come educational institutions will play an ever growing part in the socialization process.

* Division of Child Psychology, Washington University School of Medicine, St. Louis 10, Missouri.

¹ This paper is based upon an M.A. thesis submitted by AWB to the Graduate Board of Washington University.

CHILD DEVELOPMENT

For these reasons it is highly desirable that empirical study of relations between the operations of such institutions and child development be pursued with vigor. Accordingly, the present study directed itself to an investigation of relations between some aspects of children's adjustment in kindergarten and their prior attendance at a "progressive" nursery school—a problem area wherein previous researches, using a wide variety of procedures, have yielded results which are, as a body, equivocal to say the least (1, 3, 4, 5, 10, 13, 14, 19, 21).

METHOD

Design of the Study

The basic structure of the study followed a paired replicates model. Each pair consisted of two children drawn from the same kindergarten class, one of whom had attended nursery school for at least one year and the other of whom had not. Each pair was matched for social status, sex, ordinal position in the family, and IQ.

Social status was determined according to Hollingshead's two factor index (12). This index yields five status levels based upon both the education and the occupation of the principal wage earner, occupation being weighted more heavily.

Ordinal position matching was in terms of "only child," "oldest child," "youngest child," or "intermediate position." It should be noted that, in view of the relatively small size of nursery school families, matching for ordinal position provides a partial control for differences in family size. IQ was determined by administering individually to each child the Ammons Picture Vocabulary Test, Form A (2).

Finally, it should be pointed out that each pair was closely equated for age automatically as a result of grade and IQ matching.

Subjects

A total of 84 children served in this study. Half of these had attended the Washington University Nursery School during the year preceding the study (1957-1958). There were a total of 46 children in this nursery school class, but three had moved out of town and a fourth child's nonparticipation will be discussed later. All were in kindergarten when the study was made.

The children attended a total of 17 public and two private schools, chiefly in suburban communities. They were children mostly of upper-middle and lower-upper class families. Using Hollingshead's index, 64 per cent of the children fell in the highest class level, 21.5 per cent in the second, 12 per cent in the third, 2.5 per cent in the fourth, and none in the lowest. The skewed distribution obtained was, of course, a consequence of the characteristics of the nursery families, most of which were in the higher status ranges. Family sizes ranged from one to five children with a mean of 2.9. Of the 84 children who served as Ss, 60 were boys. The youngest

child was 5 years, 4 months of age, and the oldest was 6 years, 6 months. IQs ranged from 97 to 146.

Materials

Four independent graphic rating scales were constructed according to criteria suggested by Guilford (9). Each of these was separately mimeographed on 8½ by 11 inch sheets and consisted of a continuous vertical line five inches long with points marked "very good," "good," "fair," and "poor" spaced evenly along it. Equal space was left between the lowest point marked ("poor") and the origin of the scale as well as between the highest point ("very good") and the terminus.

This procedure allows the rater to choose any point along the scale which he deems appropriate (the marked points representing orientation guides) and was thus consistent with a conception of adjustment as evaluatively determined and continuous. It has the additional merit of "spreading" the ratings, thus helping to control any "central tendency" error.

The four scales produced in this manner constitute the adjustment scales. The first of these (the "activities" scale) related to the child's adjustment to kindergarten activities. It was intended to tap the general level of adjustment exhibited by the child in his routine attitudes toward and participation in usual classroom activities. A second or "group" scale had to do with the child's adjustment to his peers as this was manifest in the classroom and as perceived by his teacher. The scale referred to the general level of adjustment exhibited in the child's relations with the other children in his classroom. The third scale ("authority") had to do with the general level of adjustment exhibited by the child in his routine relations with his teacher. Finally, the fourth scale ("personal") referred to the child's personal or "inner" adjustment. It directed the rater to consider both her knowledge of the child within the classroom and outside of it and to think in terms of the general level of adjustment felt to characterize the child "as a person." These four scales were construed as dealing with at least partially independent facets indicative of general social adjustment. They reflect what may be thought of as four primary aspects of the child's general or over-all school adjustment.

A fifth rating, of relative brightness (intellectual), was obtained via a forced choice procedure. To obtain this rating, names of paired children were listed and the rater asked simply to encircle the name of that child she felt to be "brighter." The rater was instructed always to choose one, even if the two seemed about equally bright. This scale was also mimeographed on a separate 8½ by 11 inch sheet and included, in addition to the names of the critical pairs, paired names of other children selected at random from the same class. In this way it was hoped to exercise a small measure of control over possible biasing effects attributable to the rater's guessing the purpose of the ratings. The five rating scales were bound together with a face sheet containing instructions to form a six-page booklet.

Procedure

Information concerning the major wage earner's occupation and education, number of siblings, school attended, and other pertinent information was obtained by direct telephone contact with the parents of the nursery children. The same information for the control children was obtained from the records of the particular school the child was attending.

After a match had been found for a particular nursery child, the teacher was provided with a rating booklet and asked to fill it out. She remained ignorant of the purpose of the ratings until after she had completed them. The rater was left relatively free to employ whatever criteria she deemed relevant in making her ratings—only general orienting suggestions were made. Twenty-three teachers were involved in this study.²

It was not possible in all cases to find a perfect match for each nursery child from among children in the same class. Since the primary dimension for matching involved the use of a classmate who had not attended nursery school, one nursery child had to be discarded from the sample owing to the absence of any child from the class who had not attended an organized nursery school.

Children who had attended parent cooperatives were included in the study as control children since this type of organization is not designed to "teach" as is the nursery school. In a few cases it was impossible to match ordinal position exactly so it was as closely equated as possible, that is, to the next category. IQ was matched within 10 points except for one exceptionally high child ($IQ = 146$) whose closest equal was 18 points lower. Social status and sex were matched in all cases. Generally, each pair was well matched in all respects.

RESULTS

The Adjustment Ratings

After the various ratings had been obtained according to the procedure outlined above, each of the four adjustment scales was divided into 10 equal intervals and quantified by assigning to each interval a numerical value ranging from 1 for the lowest interval to 10 for the highest. In this way each rating could be scored from 1 to 10, with a score of 1 indicating the

² Unfortunately, it was not possible to obtain direct estimates of the reliability of the teachers' ratings. However, it would seem a reasonably safe assumption that reliability would be high as all of the teachers were rating children well known to them and along dimensions not at all alien to their thinking about these children. Furthermore, since a relatively large number of teachers was used, it is likely that personal peculiarities of any particular teacher which might attenuate the reliability of her ratings would be well compensated in the over-all results. Finally, since the ratings were obtained late in the school year, it is unlikely that marked changes would occur in the children themselves sufficient to alter the results presented. It might be added that in other similar (though as yet unpublished) studies these same scales have produced highly reliable ratings from preschool teachers.

poorest level of adjustment and a score of 10 indicating the highest level. Actually (as might be expected) there were no ratings in either of the highest or lowest intervals. The adjustment ratings on each scale ranged from 2 through 9 inclusive, with a modal rating of 6. Some skewness toward the low end of the scale was evident in the distribution of individual ratings.

Differences in rated adjustment between the nursery and control children were evaluated by performing independent *t* tests of the difference between the mean adjustment scores for each group on each of the four adjustment scales. The results of these analyses, which followed a procedure for paired observations suggested by Dixon and Massey (6), are reported in Table 1. In this table *N* represents number of pairs, not number of subjects.

TABLE 1

MEAN SCORES ON FOUR TYPES OF SOCIAL ADJUSTMENT IN NURSERY AND NONNURSERY CHILDREN (*N* = 42)

<i>Adjustment Scale</i>	MEAN SCORE		
	<i>Nursery</i>	<i>Control</i>	<i>t</i>
Activities	5.9	6.7	2.80**
Group	5.7	6.3	2.50*
Authority	6.3	6.5	.93
Personal	5.7	6.5	2.24*

* $p \leq .05$.

** $p \leq .01$.

As can be seen from this table, the control children receive higher mean ratings on each of the scales. The differences between the means are significant for three of the four ratings (activities, group and personal), while the difference on the fourth scale (authority) is consistent in direction but is of insufficient magnitude for statistical significance. These results indicate, then, that, with respect to personal adjustment, participation in group activities, and relations with other children, nonnursery school children are rated by their teachers as better adjusted than those who have attended nursery school. In the case of the adjustment of the child to the teacher, the difference in ratings is not significant.

The Brightness Ratings

Evaluation of the teacher's ratings of the relative "brightness" of the nursery and control children was effected by a nonparametric sign test. Because of the method by which these ratings were obtained, it seemed best to employ a method relatively free from parametric assumptions concerning the variable being measured. Further, since directional differences

CHILD DEVELOPMENT

only were being considered, the sign test for paired replicates seemed fully appropriate.

It was found that in 24 of the 42 paired cases the children who attended nursery school were rated as brighter by their teacher. By sign test this difference is not significant, supporting the conclusion that kindergarten children who have been to nursery school do not differ in *rated* intelligence from those who have not.

Adjustment and Intelligence

To determine the extent to which the various adjustment ratings covaried with the teachers' ratings of the children's brightness, point biserial correlations were calculated between the brightness ratings and each of the four adjustment ratings. For the purpose of this analysis, the children were divided regardless of nursery school attendance into two categories, one consisting of those children the teacher rated as "brighter" and the other including the children she did not so rate. Point biserial *r*s were then computed between the dichotomous ratings and each of the four adjustment scales. The resulting correlations are presented in Table 2.

TABLE 2
POINT BISERIAL CORRELATION BETWEEN INTELLIGENCE
AND FOUR ADJUSTMENT RATINGS (N = 42)*

<i>Scale</i>	<i>r</i>
Activities22
Authority13
Group27
Personal19

* $r \geq .22$ for significance at $p = .05$.

It can be seen from this table that only two of the four correlations are significantly different from zero and that both of these indicate, at best, only a slight relation between the brightness ratings made by the kindergarten teacher and her ratings on the four adjustment scales. It must be emphasized here that this finding has no relevance to any possible relations between intelligence and adjustment. Any such correlation in the present study would necessarily approach zero as a consequence of the matching procedure which may be presumed to have controlled effectively for any correlation between intelligence and adjustment. Rather, the point of the analysis performed was simply to determine whether children rated "brighter" would exhibit correlative variations in rated adjustment. From the results reported in Table 2 it may be concluded that such seems not to be the case.

As a further check on the relations between the adjustment ratings and the teachers' ratings of brightness, median tests were performed (17). No significant relation was found between brightness and any of the four adjustment ratings, as might be expected from the results of the biserial correlations.

Hence, it would appear that, when actual variations in IQ are controlled, the teachers' ratings of adjustment vary independently of their perceptions of the relative "brightness" of the children rated.

Relations Among the Scales

To determine the degree of relation among the scales, product-moment correlations were computed between the several adjustment ratings. Each type of rating was correlated with each other yielding a matrix of six correlations as reported in Table 3. In calculating each of these correlations, ratings for all 84 children in the sample were pooled regardless of whether they had been to nursery school.

TABLE 3
PRODUCT-MOMENT CORRELATIONS BETWEEN THE FOUR
ADJUSTMENT RATINGS ($N = 84$)*

Scale	Activities	Authority	Group
Authority54
Group75	.77	..
Personal55	.48	.52

* $r \geq .28$ for significance at $p = .01$.

It may be seen from Table 3 that relations ranging from moderate to strong exist between the teacher's ratings of the various facets of over-all adjustment selected for study. Since the teacher in every instance is rating "adjustment" and is furthermore rating the adjustment of a single child, it would be expected that substantial correlations would exist among these ratings. It may also be inferred from the obtained correlations that kindergarten children tend to exhibit a considerable consistency in the various aspects of their personal and social adjustment, at any rate as these are evaluated here. This finding is, of course, not especially surprising.

Finally, the fact that the correlations reported in Table 3 are rather less than unity (the mean r is approximately .60) is consistent with the assumption made in the construction of the adjustment rating scales, i.e., that each of the scales is tapping an aspect of general adjustment at least partially independent of the others. Therefore, it would appear reasonable to consider the results of the previous analyses in these terms.

CHILD DEVELOPMENT

DISCUSSION

The results of the present study quite clearly fail to support the hypothesis that nursery school attendance will enhance later school adjustment, at least relative to comparable nonattending children. On the contrary, the indications are that children who have not been to nursery school are perceived by their teachers to be generally better adjusted than are children who have attended such schools.

It warrants emphasis that this study does not demonstrate "absolute" differences in adjustment, excepting insofar as these are reflected in the way the teacher perceives the child. At any rate, it should be clear that other kinds of observers and/or different criteria might yield different results. If one assumes (as in the case in this paper) that adjustment is determined by evaluative operations of observers, it follows that characteristics *both* of the evaluated individual and of the evaluator will be relevant. In other words, assessments of the adjustment of individuals will vary as some function of the characteristics of those individuals in combination with characteristics (including criteria) of the individuals doing the evaluating. As a result, from the same raw data, two observers could arrive at different judgments concerning the "adjustment" of a person. In fact, Tindall (18) has demonstrated that remarkably low correlations exist among different criteria of adjustment.

These points may, in part, account for some of the disparate results reported in the literature. For example, some studies have employed behavior check lists which are completed by the teacher and the results of which are then pooled in some fashion to provide an index of general adjustment. In effect, the investigator provides the adjustment criteria according to his conceptions of the requisites of "good adjustment" and may, therefore, arrive at a judgment different from that which would have been made by the teacher had he or she rated "adjustment" rather than particular characteristics.

As a related point, it might be noted that a judgment of adjustment obtained by simply summing a number of segmental attributes may be quite different from a judgment which takes account of the *gestalt* of such attributes in a given individual. Taken individually, a particular collection of traits may well suggest "good adjustment," but when combined in a given individual it may not be optimal for "good adjustment." Thus, a global rating may yield very different results from a simple sum of attributes.

All of this suggests that a highly significant line of future research pointing toward clarification of some of these issues will address itself to investigation of the factors which enter into the kinds of evaluative operations alluded to above. Until a great deal more information is available, it would seem to be a more meaningful approach to deal directly with the results of these judgments for such purposes as the problem investigated

here and to partition the problem into two parts: the evaluative judgments themselves, on the one hand, and the factors which determine them, on the other.

The brightness ratings obtained here are broadly consistent with Page's (16) and other IQ studies. From the failure to find any significant differences it may be concluded that the teacher's perceptions of a child's "brightness" and her perceptions of the child's adjustment are independent dimensions.

While it is not at all simple to account for the results obtained in this investigation, we might note some of the seemingly more plausible hypotheses. Firstly, it is generally true that nursery school offers to the child a freedom and lack of restriction which may run counter to the somewhat greater conformity demanded in the kindergarten. The use of individual initiative seems to be stressed in many nursery schools whereas in kindergarten there are more directed types of activity. Those behavior patterns acquired in nursery school could influence the kindergarten teacher's perception of the child's adjustment.

It is also possible that redundancy of kindergarten and nursery school activities may lead to some primitivation of behavior. A child may become a problem in kindergarten, who has not been one before, because of the repetition of games, songs, naps, art, rhythms, etc. He may show problematic behavior because he is sincerely bored or because this previous knowledge of the curriculum may make him want to show his classmates that he knows all about it and variously try to enhance his status. This interpretation is consistent with Allen and Masling's (1) failure to find significant differences until second grade, differences which then were in favor of nursery school children. A child may take several years to "outgrow" this sort of nursery school-related difficulty, but by the time he reaches second grade the situation may reverse itself and actually become manifest in the form of better adjustment. It would naturally be the brighter children who would be affected most adversely by redundancy, and, since the mean IQ of the present nursery group was 118, it may well be that any positive effects of nursery school would be delayed in demonstrating themselves until there was no longer any repetition in the curriculum.

The rating of adjustment to the teacher showing no difference between nursery and nonnursery children is consistent with these notions. Even though the nursery child may be poorly adjusted for the first few years following nursery school, his relations with the teacher need have nothing to do with his personal adjustment or with his adaptation to kindergarten. He may still manifest friendly relations with the teacher and at the same time be bored with the repetitive activities.

Then too, as a related point, nursery school may foster development of habit patterns, not necessarily bad in the long run, but which may in a young child lead to behavior resulting in perception by the teacher of poor

adjustment. The child may not "actually" be more poorly adjusted but may strike the teacher that way because of his seeming nonchalance and/or independence. This may be why the teacher rated nursery children lower in group, personal, and activity adjustment. The child may regain his former interest and activity in new things a little later on in school. For example, Ezekiel (7) has noted egocentricity to increase with length of stay in a nursery school. Andrus and Horowitz (3) found that in six out of seven schools the correlation between insecurity and length of time spent in a nursery school was positive as measured by adult raters. Jersild and Mackey (13) note an increase in children's fights and quarrels with nursery school attendance. It may be habits such as these on which the teacher is basing her opinion. Of course, these factors may indicate poor adjustment in something like a clinical sense, but there are other things to be considered.

Apart from the kinds of issues already discussed, it is highly probable that a selective factor influences the enrollments of nursery schools. A study of 100 applicants from the waiting list of a public school nursery school in a midwestern city showed that 51 of the parents believed their children needed supervised play. Eighteen thought the nursery school educational program would be good for the child. Fifteen were employed. Seven said they had an only child and wanted playmates for him. Other reasons were "middle child trouble," "lack of adequate play space," "inability to get along socially," "his older sister enjoyed nursery school," and "we want to broaden our knowledge of child rearing" (15).

These reasons all sound very good, but there probably are others which the parents would not put down on the applications, such as: poor marital situations, parents who feel they are failing the child emotionally, and parents who feel it is a baby-sitting center while they pursue their own interests. A higher proportion of problem children (and/or children from "problem families") probably find their way to nursery school as it is set up at present. Parents encounter difficulty managing such children at home and feel that maybe nursery school is the answer, since one of the functions of nursery school is a counseling service. Some mothers may just want to dispose of the problem child for a few hours a day to settle their own nerves, without any regard for the therapeutic value of a nursery school. Voas (20), in fact, finds more emotional problems and remedial academic difficulties in children who have attended nursery school than in children who have not. It may be that nursery school did not make them problems but that they already were before starting preschool. It is even possible that such children are better adjusted than they would have been had they not attended school.

It is obviously not possible at present to choose among these alternative hypotheses or others which may occur to the reader. At any rate, it is clear that numerous research problems remain to be dealt with in this provocative area.

SUMMARY

This study compared teachers' ratings of social adjustment in 84 kindergarten children equally divided between those who had nursery school training and those who had not. Each nursery child was matched with a nonnursery child from the same kindergarten for sex, IQ, socioeconomic status, and ordinal position in the family. The appropriate teacher was then asked to rate both children separately on their adjustment to group activities, to the teacher, to other children, and on over-all personal adjustment. In addition, a brightness rating was included which asked the teacher to mark which child of the pair she felt to be "brighter." The teacher was kept ignorant of the purpose of the study.

The basic results were as follows:

1. Nonnursery children were perceived by their teacher to be better adjusted than nursery school children in personal adjustment, relations with other children, and participation in group activities. Adjustment of the child to the teacher did not produce significant differences in rating, although the direction was consistent.
2. Children who had attended nursery school did not differ in intelligence, as rated by their teachers, from those who had not.
3. The teacher's ratings of a child's adjustment varied independently of her perceptions of the relative "brightness" of the children rated.
4. Intercorrelations among the four adjustment scales ranged from moderate to high, indicating, on the one hand, a high measure of consistency in these ratings and, on the other hand, a measure of independence.

Various interpretations of these results are offered along with suggestions for further research designed to clarify some questions raised.

REFERENCES

1. ALLEN, G., & MASLING, J. An evaluation of the effects of nursery school training on children in the kindergarten, first and second grades. *J. educ. Res.*, 1957, 51, 285-296.
2. AMMONS, R. B., & AMMONS, H. S. *Full-Range Picture Vocabulary Test*. Psychological Test Specialists, 1948.
3. ANDRUS, R., & HOROWITZ, E. L. The effect of nursery school training: insecurity feelings. *Child Developm.*, 1938, 9, 169-174.
4. BONNEY, M., & NICHOLSON, E. L. Comparative social adjustments of elementary school pupils with and without preschool training. *Child Developm.*, 1958, 29, 125-133.
5. CUSHING, H. M. A tentative report on the influence of nursery school training upon kindergarten adjustments as reported by kindergarten teachers. *Child Developm.*, 1934, 5, 304-314.

CHILD DEVELOPMENT

6. DIXON, W. J., & MASSEY, F. J. *Introduction to statistical analysis*. McGraw-Hill, 1957.
7. EZEKIEL, L. F. Changes in egocentricity of nursery school children. *Child Develpm.*, 1931, 2, 74-75.
8. FISHER, M. L. Measured differences between problem and nonproblem children in a public school system. *J. educ. Sociol.*, 1934, 7, 353-364.
9. GUILFORD, J. P. *Psychometric methods*. McGraw-Hill, 1954.
10. HATTWICK, D. The influence of nursery school attendance upon the behavior and personality of the preschool child. *J. exp. Educ.*, 1936, 5, 180-190.
11. HAVIGHURST, R. J., & NEUGARTEN, B. *Society and education*. Allyn-Bacon, 1957.
12. HOLLINGSHEAD, A. B. *Two factor index of social position*. Yale Univer. Press, 1957.
13. JERSILD, A., & FITE, M. Children's social adjustments in nursery school. *J. exp. Educ.*, 1937, 6, 161-166.
14. KAWIN, E., & HOEFER, C. *A comparative study of a nursery school versus a nonnursery school*. Univer. of Chicago Press, 1931.
15. MOUSTAKAS, C., & BERSON, M. P. *The nursery school child care center*. Whiteside, 1955.
16. PAGE, J. D. Effect of nursery school attendance upon subsequent IQ. *J. Psychol.*, 1946, 10, 221-230.
17. SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. McGraw-Hill, 1956.
18. TINDALL, R. H. Relationships among measures of adjustment. In L. Gorlow & W. Katkovsky (Eds.), *Reading in the psychology of adjustment*. McGraw-Hill, 1959. Pp. 107-115
19. VANALSTYNE, D., & HATTWICK, L. A. A follow up study of nursery school children. *Child Develpm.*, 1939, 10, 43-72.
20. VOAS, W. H. Does attendance at the Winnetka Nursery tend to raise the IQ? *Yearb, nat. Soc. Stud. Educ.*, 1940, 39(II), 363-377.
21. WALTHER, J. The influence of nursery school education upon behavior maturity. *J. exp. Educ.*, 1939, 8, 164-165.

PREDICTION OF SCHOOLAGE INTELLIGENCE FROM INFANT TESTS¹

SIBYLLE K. ESCALONA* and ALICE MORIARTY

Albert Einstein College of Medicine

When applied at age levels below 18 months, the term "intelligence test" is misleading. At the preverbal stages of development the group of functions designated as intelligence, and measured by intelligence tests applicable at schoolage and beyond, has not yet emerged. The true relationship between that which is measured by infant tests and that which we later call intelligence remains largely unknown. What little we do know, however, suggests that infant IQs or DQs do not reflect functions in direct continuum with the group of functions later assessed by intelligence tests. Or, as Bayley put it in a fairly recent review of the problem, "It is now well established that we cannot predict later intelligence from the scores on tests made in infancy" (2, p. 806). With one important exception to be mentioned below, no significant correlations have been reported between test scores obtained at less than 1 year of age and intelligence test scores obtained in middle childhood and beyond.

Curiously, these consistently negative results have had an impact upon social practice (especially in the field of adoptive placement), but they have not diminished scientific concern with the area of infant testing. On the contrary, the most recent decade brought a revival of interest, as attested by a larger number of publications and by the increased time allotted to the topic at many professional and scientific meetings. In large measure the

* Department of Psychiatry, Albert Einstein College of Medicine, Yeshiva University, Eastchester Road and Morris Park Avenue, New York 61, New York.

¹ Originally presented as a paper at Biennial Meeting, Society for Research in Child Development, Pennsylvania State University, March, 1961.

CHILD DEVELOPMENT

current interest in infant testing is an outgrowth of the general interest in very early phases of human development. However, in addition many clinicians experienced in infant testing and knowledgeable about development as a whole continue to believe that infant tests elicit behavior closely related to central developmental processes. Therefore, they also consider that differences in the test performance of infants must refer to characteristics of the organism which ought to affect learning potentials and hence later intelligence test scores. And indeed for Gesell DQs obtained at 40 weeks, Knobloch, Pasamanick, and their associates have consistently reported significant correlations with Gesell scores at 3 years and with Stanford-Binet scores at that and somewhat later ages (4, 5). Two conclusions have been agreed upon by all investigators: (a) Infant scores obtained during the first half year of life do not predict later intelligence. (b) Infant tests are relatively most effective in discriminating between those who will later show neurological and intellectual deficit and those who will not, whereas they are less effective in predicting intelligence level within the normal and superior ranges.

We count ourselves among those investigators who consider that infant tests are the best available method for the systematic assessment of developmental status and for the measurement of mental growth. Our special interest is focused on very early developmental processes and their relationship to more stable patterns of mental functioning at later ages. We were therefore motivated toward further study of the possible predictive significance of infant test scores in precisely the age range where correlations with later IQs are known to be insignificant. Similarly we wished to learn more about the relationship between individual differences in test scores manifested during the first half year of life and within the normal range and comparable differences in IQ scores obtained at an age when the IQ has become as relatively stable as it ever will be.

Before turning to the empirical findings, it will be useful to turn attention away from the usual criteria of test validation and test reliability (all of which presuppose either the measurement of existing functions or the measurement of a straightforward antecedent-consequent relationship). Instead, we would call attention to currently accepted theory about basic processes of mental development.

Most of us do not view intelligence as a fixed potential or a trait, latently present in early infancy and gradually unfolding with maturation. Certainly the authors are not alone with their assumption that intelligence must be viewed as an interaction phenomenon. With apologies to Piaget's and other developmental theories we shall crudely describe intelligence as the result of a continuous stream of transactions between the organism and the surrounding field, which field is the sum total of physical and social environmental conditions. Since the organism brings to each adaptive act certain properties, both those of an intrinsic biological nature and those which are the result of the impact of previous experience upon the organism, it is

correct to say that intelligence development is at all times dependent upon what the organism is like. Yet, since it requires environmental circumstance to mobilize the organism, and since the kind of transaction which develops depends on the objective content of that to which the organism must adapt, it is equally true to say that the development of intelligence depends at all times on the experience encountered by the growing child. If Gordon Allen's use of the term "genotype" is modified so as to include properties of the organism that may have been acquired pre- or paranatally, our view approximates his in emphasizing that the phenotypic expression of the same genotypical characteristic will vary enormously, depending upon the matrix of developmentally relevant outer circumstance (1).

If all of this is reasonable, it follows that a manifest deficit in sensorimotor functioning should enter each adaptive transaction and should limit or distort the development of intelligence. It would further follow that exceptionally good organismic equipment, as manifested by acceleration or superiority of sensorimotor function, should equally enter all transactions and enable the organism to derive maximal benefit from available stimulation and thus show superior intelligence later. Such a view is compatible with the expectation that average and superior endowment may be suppressed by limiting or actively traumatic life experience—as indeed we know can be the case. It is also compatible with the expectation that unusually facilitating experience can enhance the emerging intelligence at any level of initial endowment. But our view also requires that such facilitation occurs within biologically determined limits. No amount of facilitating experience can be expected to produce a level of intelligence functioning that transcends the limits set by biological givens. In the absence of unusual limitation and of very unusual facilitation, under what Hartmann (3) termed the "average expectable environment," marked differences in the level of sensorimotor functioning should then correspond to marked differences in later intelligence functioning.

The group of children on whom we obtained infant test scores some 10 years ago, and who were retested at early schoolage, come close to meeting this last requirement. Our subjects were originally selected for "normalcy" by medical, social, and developmental criteria.² They have grown up in a midwestern town which provides a relatively benign and consistent milieu.

Of the original sample of 128 infants so selected, 70 were subsequently tested by my collaborator and by one other psychologist in connection with two separate research projects.³ In investigating the degree to which infant test findings bear a relationship to later intelligence, we were anxious to use only test scores obtained under optimal circumstances. This led to

² The sample is selected so as to include children who met criteria of normalcy; it is not necessarily a representative sample.

³ We are indebted to Dr. Herman A. Witkin, Professor of Psychology of State University of New York, for permission to use intelligence test results obtained as a part of his research on the origins of different perceptual styles.

CHILD DEVELOPMENT

considerable shrinkage in the sample, as we discarded all instances where test reports expressed doubts of the validity of test results. Thus, we shall report on a population of 58 children who were first tested between the ages of 3 weeks and 33 weeks and tested again between the ages of 6 years, 2 months, and 9 years, 3 months. The infant tests were distributed fairly regularly between the ages of 4 weeks and 32 weeks, but the great majority of retests used were administered between 7 and 8½ years. Both the Cattell Infant Intelligence Scale and the Gesell Developmental Schedule were given to each infant subject. In addition, for each infant, we had available an evaluation of developmental status based on both test protocols jointly. This "clinical appraisal" took into account qualitative factors as well (such as which areas of functioning were relatively advanced or lagging and "good failures" or minimal successes) and was a part of the *original written protocol*. It proved that these clinical appraisals could be treated as ratings, distinguishing between low average, average, high average, and superior, as judged by the total infant performance. For the schoolage assessment we used only the Wechsler Intelligence Scale for Children (WISC), although for about half the subjects Stanford-Binets and other tests are also available.

Figure 1 describes the sample in terms of intelligence ranges, as determined by each of the three tests used. Note that in order to make meaningful discriminations in this sample, we departed from convention and called "low average" the range from 90 to 99 IQ or DQ points. At the time of initial testing two DQ scores between 86 and 90 were obtained, but no IQ scores below 90 were obtained at either time. The remarkable distribution of Cattell scores (58 per cent superior) reflects a statistical attribute of this instrument, which assigns a weight of .2 to each of five items so that a 4-weeks-old who passes all items at the adjacent level obtains an IQ of 200. The Gesell DQ scores come closer to the distribution as defined by WISC IQs.

In line with what was said earlier, our expectation of infant tests was *not* that they should predict later IQs. Instead, we wished to learn what portion of the variance they may contribute to later IQ scores. Or, more accurately, we should like to learn whether assessment of developmental status in infancy significantly defines the *range* of intelligence that may be expected to emerge.

From this vantage point the most common techniques of assessing the predictive effectiveness of tests become inappropriate. If nonetheless we report a few of the parametric measures, it is for the sake of comparison with other studies. (As expected, the product-moment correlation between Cattell IQs and WISC IQs is totally insignificant ($-.045$), and the same is true for the correlation between DQ scores and later IQs ($.077$)).

As will be shown later, it proved that in our sample a relationship between infant and schoolage scores could be demonstrated only for subjects whose infant scores were obtained at ages 20 weeks and above. That is, no matter how one modifies one's expectations, those infant tests given to

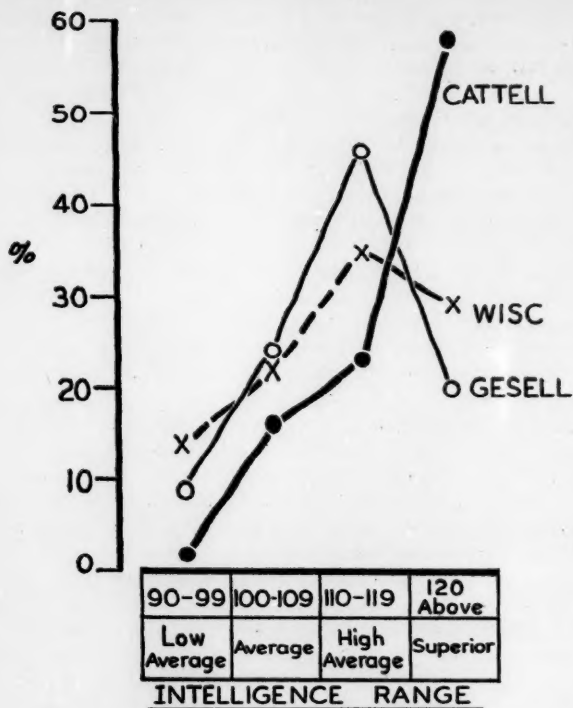


FIGURE 1—Distribution of sample over intelligence ranges in percentages as assessed by Cattell IQ, Gesell DQ, and WISC IQ.

normally progressing babies between the ages of 3 and 17 weeks bore a purely coincidental relationship to WISC IQs (except, of course, for the fact that not one of these 27 subjects obtained a WISC below 91). We therefore obtained correlations for the 31 subjects who were originally tested at 20 weeks or above and found that for the Cattell IQs $r = .212$, whereas for the Gesell DQ $r = .253$. Neither of these values is of interest. One other among the conventional measures may be mentioned, again primarily for comparison with other published values. A Spearman rank-order correlation using Gesell DQs and WISC IQs proved completely insignificant for the total population ($r_s = .062$), and very nearly so for the 31 subjects tested between 20 and 33 weeks ($r_s = .20$, which is at the .10 level of significance).

CHILD DEVELOPMENT

In these terms then, our group of children behaved very similarly to what has been reported for many other small samples of normal children. In order to test our expectation that infant tests do *something* to define the limits of later intelligence functioning, it was necessary to use a measure which would rank subjects in groups according to intelligence range. The Kruskal-Wallis one-way analysis of variance by ranks proved a suitable tool. It allows the classification of subjects into grouped rankings based on infant scores and works with the WISC raw scores for the second measure.

This test was applied to the total population in three different ways. When subjects were ranked by Cattell IQs against the WISC scores, no significant association could be demonstrated ($H = 1.73$, $df = 2$, $N = 56$,⁴ $.50 > p > .30$). When the Gesell DQ scores were used to rank our subjects, they equally failed to predict later intelligence range when the total population was used ($H = 1.3$, $df = 3$, $N = 58$, $.80 > p > .70$). Lastly, the total population was ranked in terms of the clinical appraisal which, as will be remembered, was a judgment made at the time of the original testing and based on the infant's total test performance on both instruments, with attention to qualitative features.⁵ For the total population it proved that a modest but demonstrable relationship does exist between these clinical appraisals and later intelligence range. The significance level (of an $H = 4.11$ with $df = 2$) was between .10 and .20, actually very close to the .10 level. Such as it is, the statistic is somewhat misleading in that the association between infant appraisal and later IQ is contributed almost entirely by those subjects who were tested at the age of 20 weeks and above.

On the basis of our theoretical orientation and of clinical experience we believed that—with the exception of neurologic deficit which here does not apply—infant tests would have little bearing on later intelligence when limited to a single test during the first four months of life. We had guessed that 16 weeks might be the cut-off point, for it is at this age that such more cortically determined functions as voluntary movement emerge with clarity. It proved, however, that for this sample the critical age was 20 weeks. A great many different calculations (which were hardly necessary since simple inspection left little doubt) proved that, for the 27 subjects between 3 and 17 weeks at the time of first testing, chance alone accounts for the later distribution of WISC IQs.

The Kruskal-Wallis rank-order procedure was applied to the small sample of children 20 weeks and above. When the Cattell scores were used, it was now possible to demonstrate a modest relationship as the p value fell between the .20 and .10 level ($H = 3.91$, $df = 2$, $N = 30$). We ex-

⁴ Only two subjects fell into the low average group by the Cattell criterion; these were omitted from the calculation.

⁵ These clinical appraisals do not refer to those cases where it was judged that test performance was for some reason unrepresentative of the child's developmental status. All such instances were excluded to begin with and were the major reason for the reduction in sample size from 70 to 58.

pected the Gesell Schedules to be the better instrument, chiefly because it taps a larger variety of relevant sensorimotor functions. On the other hand we consider the DQ an unfortunate numerical expression as it bears an ambiguous relationship to an irregular number of items at each age level. It is perhaps in keeping with these different attributes of the same test that, as reflected by the Kruskal-Wallis, DQ rankings showed a slightly closer relationship to later WISC scores than did Cattell IQs in that the p value fell at exactly the .10 level ($H = 6.27$, $df = 3$, $N = 31$).

When clinical appraisals were ranked against later WISC scores, and the sample was limited to subjects who were no younger than 20 weeks of age, a significant degree of correspondence emerged. The p value for this calculation falls between the .01 and .02 level of significance ($H = 8.25$, $df = 2$, $N = 31$). The sample is too small to permit of generalization, and the Kruskal-Wallis technique, it should be remembered, does no more than to demonstrate the significance of the difference between different categories—in this case intelligence ranges. Within these limitations we find food for thought in the observation that a qualitative evaluation of only test performance, i.e., not of situational factors that may have influenced performance, is considerably more effective in predicting intelligence level than are test scores.

It occurred to us that the distinctions between low average, average, high average, and superior may still be finer than the discriminating powers of infant tests with respect to later IQ scores. We therefore reassigned our subjects into only two categories, average and superior. All IQs or DQs in the range from 90 to 114 were now called average, those 115 and above were called superior. With two exceptions (which were dropped from the sample) the clinical appraisals were so worded as to make it possible to distinguish those infants who had been judged as possessing markedly superior endowment from those who showed small degrees of developmental superiority. Working only with babies tested at 20 weeks and above, we wished to compare the proportion of subjects who, from infancy to school-age, remained in the same category with those who had shifted categories. With an N reduced to 29, such a comparison offers statistical embarrassment. Short of an analysis of variance (which is not applicable to the clinical appraisals) there is no way of including the actual distance between WISC scores and the predicted range in the calculation. Thus, an IQ change from 114 to 115 constitutes a shift from average to superior just as would be the case with a shift from 100 to 130. It is perhaps for this reason that the Gesell scores, when grouped only into average and superior, failed to predict the subject's placement into corresponding categories by the WISC criterion. However, when clinical appraisals were used the same simple comparison of those who stayed within the range established in infancy with those who moved into the other range proved highly significant. Twenty-three or 79 per cent of the 29 subjects remained in the same category. A binomial test of proportion applied to this result yields an error estimate at the .002 level.

CHILD DEVELOPMENT

It may be of incidental interest that in this small sample all of the six shifts occurred among the 22 subjects originally assessed as superior. Our clinical appraisals tended to err in overestimating endowment. They did not, in this sample, underestimate the range of later intelligence functioning.⁶ (Yet for the superior group alone the proportion who remained within category was significant at the .034 level).

SUMMARY AND CONCLUSIONS

Fifty-eight children selected for normalcy by medical, social, and developmental criteria were tested between the ages of 3 and 33 weeks. For each infant we obtained (a) Cattell IQ scores, (b) Gesell DQ scores, and (c) a "clinical appraisal" based on total test performance. These subjects were again tested between the ages of 6 years, 2 months, and 9 years, 3 months, at that time receiving the Wechsler Intelligence Scale for Children (WISC).

On theoretical grounds, and in view of previous findings, parametric measures of correspondence between infant scores and later intelligence test scores were not expected to yield significant values. In this respect, results confirmed our anticipation. The hypothesis under investigation was that, in the absence of major environmental variation, infant tests will predict the *general level* of later intelligence functioning within the average and superior ranges. Further, we compared the relative predictive power in terms of intelligence ranges of Cattell IQs, Gesell DQs, and a clinical appraisal of both test protocols in combination.

It proved that, when infant tests were administered to subjects younger than 20 weeks of age, no method of appraisal predicted later intelligence range. For subjects tested between 20 and 32 weeks of age inclusive, both the Cattell IQ and the Gesell DQ bore a positive relationship to later intelligence level, as assessed through a one-way analysis of variance by ranks. However, this relationship exceeds chance at the .10 level and is therefore of limited predictive value. When the same test is applied to subjects tested at 20 weeks and above but ranked on the basis of clinical appraisal, infant assessment significantly predicted differences in later intelligence range ($.02 > p > .01$). When infant assessments were examined for their ability to distinguish merely between subjects who would later be of average or of above average intelligence, clinical appraisal (but neither of the test scores) achieved this discrimination at a highly significant level (.002).

We concluded that infant tests administered under optimal conditions at 20 weeks of age and above can approximate the prediction of later intelligence range if environmental conditions are relatively stable and benign. It was further concluded that, for children in this age range, a clinical ap-

⁶ This direction of change is in accord with results reported by Knobloch *et al.* (4, 5) and can be thought of in terms of their hypothesis to the effect that detrimental environmental influences reduce intellectual achievement of which the organism is potentially capable.

praisal based on total test performance on two test instruments is a better predictor of later intelligence range than are scores obtained from either the Cattell or the Gesell tests alone.

It is suggested that research dealing with the contribution of biological and maturational factors to the development of later intelligence may become more rewarding once it is assumed that developmental examinations are not tests of intelligence. If infant tests are viewed as assessing only one among several determinants of intelligence, their usefulness for research purposes is not determined by the usual criteria of test validation, and the failure to predict later IQ scores does not relegate infant testing to the limbo of useless enterprises.

REFERENCES

1. ALLEN, G. Intellectual potential and heredity. *Science*, 1961, 133, 378-379.
2. BAYLEY, N. On the growth of intelligence. *Amer. Psychologist*, 1955, 10, 805-818.
3. HARTMANN, H. *Ego psychology and the problem of adaptation*. International Universities Press, 1958.
4. KNOBLOCH, H., & PASAMANICK, B. The distribution of intellectual potential in an infant population. In *The epidemiology of mental disorder*. American Association for the Advancement of Science, Publication No. 60, 1959. Pp. 249-272.
5. KNOBLOCH, H., & PASAMANICK, B. An evaluation of the consistency and predictive value of the 40 week Gesell Developmental Schedule. In C. Shagass & B. Pasamanick (Eds.), *Child development and child psychiatry*. Psychiatric Research Reports of the American Psychiatric Association, 1960, No. 13. Pp. 10-31.

COMMENT ON "THE RELATIONSHIP BETWEEN RIGIDITY-FLEXIBILITY IN CHILDREN AND THEIR PARENTS"

JOHN WILLIAM ASHER*

University of Pittsburgh

In the 1959 article by Blum (1) intercorrelations among mothers, fathers, and children on measures of rigidity-flexibility were presented. In addition to the three simple two-variable correlations among the three groups, multiple correlation of mothers' and fathers' rigidity with their child's score was obtained (*see* Table 2 of the cited article).

There were no significant correlations among the three variables taken two at a time. A multiple correlation of .459 was found for mothers' and fathers' scores taken as the dependent variables and the children's scores taken as the independent variable with 17 scores on each variable.

The author concludes, "These findings support the prediction, namely, that in a normal sample the child's relative rigidity-flexibility pattern is positively related to this variable in his parents. Although no significant relationships were found for the child and either one of his parents separately, a significant relationship was found when both parents were considered. Such findings should further support investigations which attempt to understand and measure the influence and effect of *both* parents, as a family entity, upon the growth and development of the child" (1, pp. 303-304).

To this reader, the significance levels presented with the size of the multiple correlation obtained, especially with the relatively small size of the sample, seemed unusual. A standard statistical table was consulted, Table D, Coefficients of Correlation and *t* Ratios Significant at the 5 Per Cent Level . . . and at the 1 Per Cent Level . . . for Varying Degrees of

* School of Education, University of Pittsburgh, Pittsburgh 13, Pennsylvania.

CHILD DEVELOPMENT

Freedom (2, p. 609). For a three-variable relationship with 17 scores on each, 14 degrees of freedom remain, $17 - 3 = 14$. Thus, the table is entered with 14 *df* and with three variables. We find that a multiple correlation of at least .590 is needed to reject the null hypothesis of no correlation at the 5 per cent level and .694 at the 1 per cent level. Obviously, the multiple correlation reported is not of this size. Even a simple correlation of this size would not be considered significantly different from zero with just 17 observations.

Thus, the report that "... a significant relationship [between child and parent's rigidity-flexibility measures] was found when both parents were considered" (1, pp. 303-304) is in error. From the data provided it must be concluded that there is *no* relationship. The further theorization derived from these data would also need to be reconsidered.

However, it is believed that the opposite conclusions from those reported should not be accepted. It would appear that the small number of observations would almost invariably lead to indeterminant results unless significant findings ensue. The results must be accepted for the moment. However, for experimenters, the viewpoint should be that an interesting area of research has just started to be investigated and needs further research with larger numbers of observations before more definitive conclusions can be reached.

These comments do not of course detract from the other findings and developments of the study.

REFERENCES

1. BLUM, A. The relationship between rigidity-flexibility in children and their parents. *Child Developm.*, 1959, 30, 297-304.
2. GUILFORD, J. P. *Fundamental statistics in psychology and education*. McGraw-Hill, 1950.

Notice to Contributors

Child Development is devoted to basic research and theory concerning the growth and development of the child.

Manuscripts should be submitted in duplicate, each copy bearing only the title as a means of identification. The title, author's name, and institutional affiliation (together with any footnotes which reveal the identity of the author) should be submitted on separate pages.

Contributors are asked to refer to the *Publication Manual of the American Psychological Association* (1957 Revision) in the preparation of manuscripts.

The following rules require special emphasis:

1. All copy, including indented matter, should be typed double-spaced on white standard paper. Lines should not exceed six inches in length.

2. Footnotes should be few in number. They should be indicated in the text by superscript arabic numerals, numbered consecutively throughout the manuscript. The footnotes themselves should be typed double-spaced on a separate page placed at the end of the manuscript.

3. Tables and figures are costly to print and should be used only when essential. The author will be asked to contribute to the cost of composition of tabular and graphic material which is excessive in amount or unduly complicated in form. Each table should be typed on a separate sheet. Insert a guide line (e.g., Table 1 about here) at the appropriate place in the manuscript. Figures should be prepared with black India ink, suitable for photography. Legends typewritten on an ordinary typewriter are unacceptable.

4. References in the text should be indicated by numbers in parentheses that refer to the list of references at the end of the article. Do not use footnotes for this purpose. In the list of references, the arrangement should be alphabetical by author's surname, or for multiple authors by the name of the senior; for any other author, by order of date. The following standard form should be observed:

Jones, H. E. Development in adolescence. D. Appleton-Century, 1943.

Shepard, W. O., & Schaeffer, M. The effect of concept knowledge on discrimination learning. Child Developm., 1956, 27, 173-178.

The World List of Abbreviations is used for journals.

The manuscript submitted should represent the final form in which the author wishes the paper to appear. All changes in proof (other than printer's errors) will be charged to the author.

All manuscripts should be addressed to William E. Martin, Editor, Child Development Publications, Purdue University, Lafayette, Indiana.

CURRENT MONOGRAPHS

of the Society for Research in Child Development, Inc.

Helen L. Koch

THE RELATION OF CERTAIN FORMAL ATTRIBUTES
OF SIBLINGS TO ATTITUDES HELD TOWARD EACH
OTHER AND TOWARD THEIR PARENTS

1960 \$3.50

Nancy E. Wood

LANGUAGE DEVELOPMENT AND LANGUAGE DIS-
ORDERS: A COMPENDIUM OF LECTURES

1960 \$2.75

Irwin M. Marcus, *et al.*

AN INTERDISCIPLINARY APPROACH TO ACCIDENT
PATTERNS IN CHILDREN

1960 \$2.50

Bernard Farber

FAMILY ORGANIZATION AND CRISIS: MAINTENANCE
OF INTEGRATION IN FAMILIES WITH A SEVERELY
MENTALLY RETARDED CHILD

1960 \$2.75

IOWA CHILD WELFARE RESEARCH STATION. STATE
UNIVERSITY OF IOWA. THE FORTIETH ANNI-
VERSARY: 1917 - 1957

1959 \$2.50

Ali A. Kawi and Benjamin Pasamanick

PRENATAL AND PARANATAL FACTORS IN THE DE-
VELOPMENT OF CHILDHOOD READING DISORDERS

1959 \$3.00

Ruth Hoeflin and Leone Kell

THE KELL-HOEFLIN INCOMPLETE SENTENCE
BLANK: YOUTH PARENT RELATIONS

1959 \$2.25

Lucille Aust Hunt

A DEVELOPMENTAL STUDY OF FACTORS RELATED
TO CHILDREN'S CLOTHING PREFERENCES

1959 \$1.75

(A mimeographed list of all monographs available in the series since 1935
will be sent upon request.)

order from

CHILD DEVELOPMENT PUBLICATIONS

Purdue University, Lafayette, Indiana

